Institute
and Faculty
of Actuaries

Impactability Modelling for Population Health Management

A review of current concepts and practices

by the IFoA PHM Working Party

4 June 2020

Members of the Population Health Working Party:

Chair: A. Shah

Deputy Chair: D. Beddows

I. Bakbergenuly, J. Buckle, C. Bull, C. Campbell, M.S. Chan, S. Culkin, K. Dreyer, M.R. Elsheemy, M. Flint, T. Hayward, S. Lorrimer, C. Martin, T. Mason, S. McKeigue, U. Moorthy, L. Morgan, J. Robertson, J. Seymour, C. Swatton, J. Umpleby.

Keywords

Population Health Management; Impactibility, Impactability, Modelling; Risk Stratification; Actuarial Analysis; National Health Service

Table of contents

Section 1. Executive Summary

1.1 Introduction

The UK's National Health Service (NHS) has identified Population Health Management (PHM) as a key component of its long-term plan to develop integrated local healthcare systems that provide the right care at the right place and at the right time. PHM uses data to inform the development of interventions tailored to local at-risk population cohorts, aiming for improved outcomes with reduced unwarranted variation between cohorts.

The actuarial skillset is well-placed to support the NHS in building and deploying the demographic, risk modelling and analytical capability required to successfully embed PHM into the formation of integrated care systems. As a regulated profession, members of the Institute and Faculty of Actuaries (IFoA) are relied upon by their stakeholders to produce robust models, analyses, judgements and insights from complex datasets, including being clear about the merits and limitations of each. The combination of analytical capability and professional rigour ensures that the NHS and other healthcare stakeholders, including patients, can have confidence in the quality and relevance of actuarial PHM recommendations. The IFoA also recognises the limited analytical capacity that currently exists across the NHS, and is ready to help build more advanced analytical capability that will be necessary to harness the potential benefits of PHM.

As an initial step, the IFoA has formed a Population Health Management Working Party in association with NHS England, to enhance the NHS's use of data and analytics in PHM and explore what value the actuarial skillset may bring to drive that change. The Working Party is comprised of actuaries, academics and health professionals, with international representation and from a wide range of organisations including the NHS.

We envisage a series of reports to be published by the PHM Working Party. This first report focuses on developing a common technical and practical understanding of impactability, a key PHM concept for which there is not yet a significant body of literature. The report is structured around addressing the following questions:

- What is impactability modelling?
- How is it currently applied in practice?
- What theoretical models exist?
- What are the ethical considerations?

Further background is provided in Section 2.

1.2 Defining impactability

Many risk stratification and segmentation models have focused historically on identifying population groups that have a high risk of experiencing an adverse event, or who have a high cost profile. However, the success of risk stratification in managing demand effectively at the whole-population level depends not just on identifying those most at risk of an adverse event, but rather in identifying those who are most at risk *and* most likely to respond positively to a given intervention, i.e. to be 'impactable'.

impactability does not have a universally accepted definition, so a key area of focus for the working party was to try to bring clarity to what impactability is. This resulted in the following working technical definitions of "impactability" and "impactability modelling":

impactability: defines the degree to which different sub-populations will benefit from a range of interventions;

and

impactability modelling: uses this information to tailor appropriate interventions within agreed boundaries for the 'value' gained from resources spent.

The choice of words and phrases used in these definitions is analysed in detail in Section 3. However, this paper does *not* attempt to define what 'value' is, since in practice this will relate to a combination of considerations that will need to vary between different applications, geographies and population cohorts.

It is important to note that impactability modelling is not the same as risk stratification, but they are closely related. Risk stratification identifies people who are most likely to experience adverse health outcomes, while impactability modelling suggests which changes to care are most likely to lead to better health outcomes. Risk stratification statistical tools are widely used in the UK; in contrast, impactability modelling has tended to be left to individual clinicians' subjective judgement, rather than being carried out in a data-driven way using statistical models.

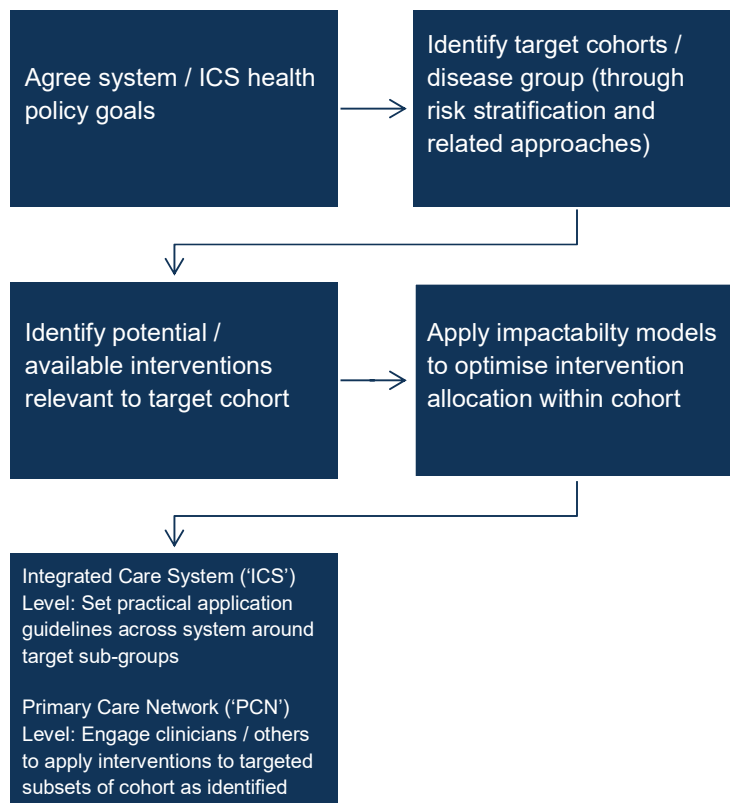The role impactability modelling can play within a wider PHM programme is summarised in Figure 1 below.

Figure 1: The role of impactability modelling within a PHM programme

1.3 Practical challenges and considerations

In order to identify the key thematic practical issues surrounding impactability modelling, we:

- completed a literature review; and
- designed and conducted a series of semi-structured interviews of a sample of ten key individuals who are involved in the development, application or study of PHM programmes involving risk stratification and impactability modelling.

The main considerations arising from the literature review and interviews are outlined below, and Section 4 presents the detailed findings. We note that there are only limited existing examples of impactability modelling being applied in practice, particularly within the NHS, and more trials would enhance the evidence base that was available for the current report.

- **Understanding of impactability modelling and its benefits/disbenefits** - There is not yet a consistent awareness and understanding of the definition of impactability modelling, meaning its use is ad-hoc with pockets of good practice. A tendency towards short-termism in the NHS needs to be addressed, given that the benefits of impactability can accrue over a longer time period. Other factors affecting the consistent understanding of impactability and its benefits include:
    - the timing of the intervention;
    - defining the appropriate population cohorts for intangible categorisations such as 'frail older people';
    - the transferability of impactability models between different geographies and/or time periods;
    - ethical challenges (see Section 6); and
    - the effect of variations in how well patients comply with their prescribed interventions.

  It is therefore important to keep the objectives of a given programme front of mind. For example, the threshold for action in a programme aimed primarily at reducing costs will differ from one aiming to maximise aggregate health status or cost-effectiveness. We hope that the focus on defining impactability in Section 3 can help drive some consistency and greater awareness of how it can help in a broader PHM programme.

- **Data** - Data is fundamental for the fitting of impactability models, and it is imperative that its limitations are well communicated and understood. The following types of data are the most relevant to impactability modelling:
    - activity and administrative data,
    - electronic healthcare records,
    - social care data,
    - randomised trial data,
    - derived or composite data (e.g. risk scores), and
    - wider datasets relevant to particular use cases (e.g. from pharmacies, opticians, education, criminal justice, etc.).

  Key data considerations include:
    - the availability of relevant and suitably linked data sources,
    - access and regulation,
    - quality, consistency, completeness and vagueness.

  We recommend engaging with data stakeholders early and setting up a clear governance structure for data handling.

- **Analysis** - There is a challenge in sourcing sufficient analytical capability to carry out and clearly communicate the outcomes of impactability modelling, in collaboration with clinicians

and patients. Analytical teams across NHS Integrated Care Systems will need to be scaled up if impactability modelling is to become mainstream. There will also need to be investment in developing analysts' communication skillset, and in providing training to end users of impactability modelling tools.

- **Organisational issues** - We identified a range of organisational factors that affect the practical application of impactability modelling. These include:
  - relationships and trust within and between teams in the NHS, social care and other organisations;
  - what PHM and impactability means to clinicians at the point of care, and how they use the analytical results to inform their subjective decision-making;
  - the inter-operability of records systems;
  - the extent to which performance measures incentivise organisations to implement the results of impactability modelling;
  - staff training; and
  - continual reform and embedding impactability into a wider data-driven change management process.

- **Evaluation** - There are currently no standard evaluation procedures to ensure that impactability modelling is achieving its intended consequences. It will therefore be necessary to agree a way to demonstrate that an intervention is working, including evidence to measure the realisation of benefits and compare them to programme costs and the opportunity costs of alternative interventions or investments. Key practical issues relating to evaluation include:
  - adopting a holistic approach encompassing all effects arising from an intervention;
  - selecting outcomes for evaluation that best enable assessment of achievement of the overall objectives of an intervention; and
  - using pilots to trial proposed initiatives and evaluation methodologies in the real world.

1.4 Potential approaches and models

There are various potential models for measuring and predicting impactability. We evaluated a number of example options providing coverage of a range of the possible approaches available. Each approach has its own advantages and disadvantages whose relative importance will vary according to the specific intervention programme and context under consideration. Our aim in this review was to provide an understanding of the range and types of approaches and models available to encourage readers to consider which approach(es) may be appropriate to their context.

We have classified each impactability model under one of the following categories:

- **Traditional models** - These are already in common usage for individual patients, albeit inconsistently applied and not necessarily referred to by users as 'impactability' models. These models are based on information held in a patient's medical records, and could in future be scaled up to apply across entire population cohorts if the associated data challenges can be overcome. In Section 5 we have evaluated an example of a traditional model that involves prioritising patients for an intervention based on their gap score, which is a measure of the gaps in the care they have received compared to that recommended by the relevant National Institute for Health and Care Excellence (NICE) guidelines.

- **Statistical models** - These apply more advanced statistical techniques to patient data primarily sourced from electronic medical records and hospital data systems, calibrated for large datasets. In Section 5 we have evaluated two examples of statistical models:
  - Condition severity benchmarking, which involves identifying individuals (or cohorts) who have more severe conditions (or higher service utilisation) than would be expected for their given risk characteristics.
  - Propensity to benefit scores, which involves prioritising high-risk individuals according to a score calculated via a statistical model that predicts the benefit of an intervention before an individual joins the intervention programme. The 'benefit' is a pre-defined, condition-specific measure linked to a given intervention.

- **Survey-based models** - These involve the collection of additional data via patient questionnaires because it is not otherwise captured on a routine basis. In practice this means that it is limited to smaller groups of patients than the other categories of model. In Section 5 we have evaluated an example of a survey-based model that involves equating impactability with 'patient activation', which is a survey-derived measure of the ability an individual has to manage their own care.

We evaluated the key advantages and disadvantages of each of the four selected example models by considering criteria relating to four evaluation themes: patient outcomes and experience; experience for healthcare professionals; healthcare utilisation; and technical concerns. Table 1 summarises the key observations arising from this evaluation exercise.

	Pros	Cons	Data Required
Gaps-In-Care Scores	Based on proven clinical theory Encourages consistency in prescribed treatments Ability to measure the effectiveness of implementation Likely to reduce health inequalities	Requires agreed list of recommended care/treatment/intervention guidelines for each condition and patient group. Challenge to set appropriate weights for each element of care.	List of clinically-proven treatments for every combination of condition and patient "type" Record of care that each patient has received All diagnoses that each patient has received
Condition Severity Benchmarking	Uses data that is readily available (electronic medical records, hospital records, community and social care provider records) If the aim is to reduce service utilisation or reduce risk or reduce condition severity, then this method creates the metrics for measuring the success of itself.	Many possibilities for defining the risk score, condition severity and service utilisation and there is not a single clear preferred choice for these metrics. Need clinical expertise to set the appropriate interventions for the range of values of each metric.	All diagnoses and health services that each patient has received All factors that support the risk score Model could be built with aggregated data only rather than individual patients.
Propensity-to-Benefit Scores	The method lends itself to statistical methods, which actuaries can deploy. Standard software can be applied for fitting the logistic regression models.	Many possibilities for defining the propensity-to-benefit score and there is not a single clear preferred choice for this metric. Questionable accuracy of results, such as false positives.	All diagnoses and health services that each patient has received All factors that support the propensity-to-benefit score Specific preventive programme data A measure of what "benefit" means for each condition and treatment
Patient-Activation Scores	Helps to reduce avoidable use of healthcare services Encourages more personalised interventions Likely to reduce health inequalities	Data not readily available across whole populations Can only be practical deployed for small segments of the population, which must be first identified by some other method. Data vulnerable to subjective responses of patients. Existing versions of scores are commercially-owned so not accessible for study by this Working Party.	Patient survey data Weightings for each element of activation/engagement need to define and quantify how important self-management is to the success of a given intervention.

Table 1: The pros, cons and data requirements of selected impactability models

1.5 Ethics, patients and the public

The concept of fairness, or health inequalities, is a particularly important ethical consideration when looking to model or predict the impact of different interventions, because it can influence whether the intervention is recommended or not. When building, choosing, parameterising and using a model there are many choices faced by the user. The decision-making process is informed by prior experiences, values and beliefs. These elements can either be consciously or unconsciously embedded into a model, and could lead to unfairness and greater health inequalities. It is therefore essential that the likely impact of a proposed course of action on health inequalities is explicitly assessed before a final decision is made.

Section 6 sets out how a deliberative consideration of ethics, including the involvement of patients and the public, can support this relatively new area of health research by guiding the resulting decision-making to be fairer and more objective.

We have split the ethical considerations into two sections:

- **The ethical considerations of the data inputs into an impactability model** - This includes being clear about the purpose and expected patient/population benefit, so that only the data that is necessary is collected, stored and used. The relevant data sharing agreements and privacy impact assessments should be completed, open data and interoperability standards should be used, and data security should be made integral to the design of the model. It is also important to be transparent about limitations and biases in the data and the intended use of data.

- **The ethical considerations when using model outputs** - This aligns to classic public or population health ethics considerations, structured around beneficence (do good), non-maleficence (do no harm), fairness, and autonomy, as well as additional principles such as utility, transparency and procedural justice. For impactability modelling, it is particularly important to:
    - Factor in real-world considerations for evaluating and implementing model outputs;
    - Actively assess fairness in the selection of interventions, and how they may affect health inequalities;
    - Incorporate an ethical framework for impactability modelling such as that proposed in Section 6; and
    - Consider whether refinements to the modelling process and/or ethical framework are appropriate for any given use case.

Section 6 includes the outline of a 'reflexivity exercise' that can help practitioners to surface and record underlying beliefs and values as a way to address the potential for model outputs to introduce bias or unintentionally increase health inequalities.

It is also important to consider how best to engage with patients and the public in relation to the data being used, how the model is designed and developed, and how the model outputs are used. Those working in this area may need to seek advice regarding compliance and appropriateness in accordance with laws and regulations that are specific to the context they operate within.

1.6 Next steps

The Working Party welcomes feedback on this initial report. In future phases of work, we are considering a range of further areas to explore including:

- Reviewing a practical application of impactability modelling in detail; and

- Considering other analytical aspects of PHM that we believe could benefit from additional research by the Working Party.

Section 2. Introduction

2.1 Background

The UK's National Health Service (NHS) is facing the twin challenges of increasing demands for healthcare and constrained funding. To address this, the NHS is adopting new models of care, developing integrated local healthcare systems to provide "properly joined-up care at the right time in the optimal care setting" (NHS England, 2019). Key to the success of this will be understanding the health and related socio-economic characteristics of the population - for example, current patterns of care demand, forecasting future demand, and predicting the effects of interventions for different groups of people. The actuarial skillset has much to offer in building and deploying such demographic, risk modelling and analytical capability.

The Institute and Faculty of Actuaries' (IFoA) Population Health Management Working Party (hereafter, the Working Party) was formed in association with NHS England to explore the use of impactability modelling for Population Health Management (PHM), and review the current concept and practices. There are 23 members of the Working Party, including actuaries, academics and health professionals, with international representation and from a wide range of organisations, including insurers, consultancies and the NHS.

2.2 PHM and impactability modelling

The King's Fund definitions of "Population Health" is:

> **Population Health is an approach aimed at improving the health of an entire population. It is about improving the physical and mental health outcomes and wellbeing of people, whilst reducing health inequalities within and across a defined population. It includes action to reduce the occurrence of ill-health, including addressing wider determinants of health, and requires working with communities and partner agencies.** (Kings' Fund, 2018)

A commonly used NHS definition of Population Health Management is:

> **Population Health Management improves population health by data driven planning and delivery of proactive care to achieve maximum impact. It includes segmentation, stratification and impactability modelling to identify local 'at risk' cohorts - and, in turn, designing and targeting interventions to prevent ill-health and to improve care and support for people with ongoing health conditions and reducing unwarranted variations in outcomes.**
> (NHS England et al., 2019)

impactability modelling is therefore just one facet of a set of PHM tools and techniques available to health commissioners to help better target interventions to populations, thereby improving overall Population Health. We propose a more precise definition of impactability and impactability modelling in Section 3.

2.3 Aim of the working party and scope of this report

The working party has a broad goal to draw on a multi-disciplinary, multi-geography team to improve the NHS's use of data and analytics in PHM. We envisage a series of reports as we complete different stages of the work required to meet that goal.

Many risk stratification and segmentation models have focused historically on identifying population groups that have a high risk of experiencing an adverse event, or who have a high cost profile. However, the success of risk stratification in managing demand effectively at the whole-population level depends not just on identifying those most at risk of an adverse event, but rather in identifying those who are most at risk and most likely to respond positively to a given intervention - i.e. to be 'impactable'. In choosing to focus on impactability in this initial report, we recognised that there is already a significant body of literature on the use of risk stratification in PHM (e.g. Lewis 2015; Hippisley-Cox & Coupland 2013) . Therefore, this report does not attempt to address the advantages and disadvantages of common population segmentation or risk stratification tools and methods, as these have been discussed at length elsewhere. Instead, this first report of the working party focuses specifically on impactability modelling by bringing together existing knowledge, establishing a solid definition, providing a guide to practical applications and setting out the theoretical foundations.

In considering the framework for this report, the working party convened four separate workstream to investigate the current knowledge and usage of impactability modelling in the UK, asking the following questions:

1. What is impactability modelling?
2. How is it currently applied in practice?
3. What theoretical models exist?
4. What are the ethical considerations?

The emphasis in the literature so far has been on understanding the current use of impactability modelling, rather than necessarily promoting a particular course of action. Nevertheless, this report will be of significant value to the practitioner in its discussion of the current relevant literature, its overview of commonly used impactability methodologies and its practical recommendations on data, the technical aspects of modelling and ethical considerations. We hope this report will be of concrete use to a variety of stakeholders - from policy-makers and commissioners/managers wanting to understand more about what impactability modelling does and does not do, to clinicians and analysts looking for a practical guide.

Section 3. Defining impactability and impactability modelling

Lead authors of Section 3: J. Buckle, Clare Campbell

impactability modelling involves identifying groups of people who would benefit most from specific treatments, interventions and support as part of a PHM strategy. impactability analysis allows a health system to find the people who are most likely to respond positively to particular changes to their care. There are specific benefits to impactability modelling as it enables care provision to be aligned to peoples' needs in a way that achieves better outcomes for limited resources. impactability modelling is not the same as risk stratification, but they are closely related. Risk stratification identifies people most likely to experience adverse health outcomes, while impactability modelling suggests which changes to care are most likely to lead to better health outcomes. Risk stratification statistical tools are widely used in the UK; in contrast, impactability modelling is often left to individual clinicians' subjective judgement, rather than being carried out in a data-driven way using statistical models.

We note that several definitions of impactability already exist in the literature (Lewis, 2015, Lewis, 2010, Duncan, 2004)[1] . The working technical definitions, as defined in this report, of impactability and impactability modelling are:

impactability: defines the degree to which different sub-populations will benefit from a range of interventions,

And;

impactability modelling: uses this information to tailor appropriate interventions within agreed boundaries for the 'value' gained from resources spent.

If we unpack these definitions, there are several key elements:

- **Different sub-populations** = groups of individuals who may be defined as a result of risk stratification and/*or* who could be identified through the impactability model. Identifying sub-populations through risk stratification would, for example, help distinguish between the three types of cohort below.

  1. Who are the individuals in the general population who might benefit most from (healthcare) interventions in reducing the risk of incidence of, for example, a long term disability or acute episode?

  2. Who are the individuals with a long term disability who might benefit most from (healthcare) interventions in increased access to recovery?

  3. Who are the individuals who have previously recovered from a previous long term disability who might benefit most from (healthcare) interventions to reduce the risk of recurrence?

- **The degree to which different sub-populations will benefit** = the impactability model will use all the factors we know about homogenous groups of individuals to determine where interventions are likely to have most impact - based either on evidence-based assumptions or from clinical experience.

---

[1] "impactability models are approaches "which aim to identify the subset of at-risk patients for whom preventive care is expected to be successful" (Lewis, 2010). And "A predictive impactability model may be defined as one that predict[s] who will acquire a disease, an adverse event related to a disease, or change from one health (functioning) state to another, where these outcomes are impactable with some specific intervention such as taking or stopping a medication, doing a test, reducing avoidable medical costs, making a behavioural change, or changing the person's environment" (Duncan, 2004).

- **Using this information to tailor appropriate interventions** = the impactability model enables you to assess interventions (from the existing range of interventions) and target these more precisely given patients' capacities to benefit.

- **Within agreed boundaries for the 'value' gained from resources spent** = where a group of patients have been considered the least impactable do we create high-cost, low-impact interventions for them? In all healthcare systems there is a practical limit to the intervention resources and these resources should be used in the most effective way.[2] We discuss the ethical considerations for impactability modelling in Section 6.

- **The 'Value' gained** = one example of an outcome; could be measured in many ways, including narrow clinical indicators, process measures, financial/economic or broad health measures.

We have assumed that we are drawing from the existing universe of interventions. We are therefore focusing on the service design aspect to make NHS-paid interventions more effective, rather than proposing entirely new interventions for reimbursement.

3.1 Principles that define impactability modelling

There are several key principles that characterise impactability modelling:

- **Cohorts:** focused on population-defined cohorts likely to be predictive of future risk of sub-optimal health outcomes, rather than just disease-defined. These could include cohorts defined by socio-economic groups or other characteristics.

- **Time Dimensions:** impactability modelling may have time frames ranging from less than one year to decades.

- **Outcomes:** impactability modelling can incorporate a range of outcomes, rather than just resource use or financial outcomes. For example, healthy life years, mortality, quality of care (at an individual and population level) and staff / patient satisfaction improvements.

- **Policy-driven:** impactability modelling facilitates the achievement of overall policy goals for the NHS, such as "equal access for equal need".

- **Distributions, rather than just the mean or average outcome:** Incorporates consideration of the distribution of outcomes, rather than just shifting the average outcome.

- **Evaluation Methodology:** There is up-front consideration of the evaluation of any intervention, along with consideration of the appropriate methodology, tools and data required to overcome statistical issues such as regression to the mean, selection bias, small sample sizes etc.

- **Ethics Dimension:** impactability modelling incorporates patient-preference or patient-willingness, as well as clinically-defined need.

---

[2] We note the NICE cost-effectiveness (Cost/QALY) broad measure as one way to indicate the relative value of different interventions (NICE, 2013)

3.2 Where can I use impactability modelling?

The kinds of problem statements that can benefit from impactability modelling can be both generic and specific - some example problem statements are listed in Table 2. The most common use case is 'we have a risk stratification tool but would like to know which of our range of preventive interventions we should offer to different individuals within the high-risk group'. The problem statements below are not exhaustive, but hopefully give the reader an idea of the kinds of issues that impactability modelling could help clarify.

Generic problem statement	Clinically specific statement
We would like to predict which sub-populations are most likely to benefit from different health and care pathways, while considering the impact on inequality of outcomes.	Should we put a specific subset of patients with hypertension or hyperlipidaemia on aspirin or not? What is their likely capacity to benefit clinically? What is the financial return on investment over a specified time period?
We would like to identify missed elements of pathways of care for different sub-populations and identify the best ways to fill those gaps.	How should we decide which patients not currently taking statins would benefit from them?
We have an intervention that we think will work better for some groups of patients than others and we want to know how to identify who we should be putting forward to receive the intervention.	If we want to offer gripped slippers to a group at risk of frailty, how do we decide where to target the slippers to those at greatest risk?
We want to know the financial outcome (financial cost-benefit) for prescribing a certain drug to one group of patients versus a group with different characteristics or risk factors.	If we have three potential care pathways for chronic depression as a co-morbidity, how can we effectively match each care pathway to the set of people most likely to benefit and understand the impact on overall clinical outcomes, as well as distribution of outcomes in each cohort and across cohorts?
We want to understand the clinical value gain overall and the distribution of the clinical value gain from targeting a set of tailored interventions to groups sub-divided according to socio-economic characteristics.	

Table 2: Examples of problems that can benefit from impactability modelling

3.3 Identification of interventions

As a starting point, the need for a PHM intervention can be identified from investigating the gaps in care. Examples might include things like failure to collect prescriptions, non-attendance at appointments or indicators of poor outcomes, such as high smoking rates, poor control of symptoms, or complication rates.

National level gap-analysis of local organisations might identify gaps in the service that need to be filled. Gap-analysis of actual versus optimal practice may identify a potential to improve outcomes by ensuring a greater proportion of the population are treated optimally. The resulting change is not necessarily at the patient level but may be at the organisational or national level by restructuring team structures, or incentive schemes. For example, making sure oncologists have easy access to best practice treatment pathways, might reduce variation in care, improve quality of care and reduce costs.

The role impactability modelling can play within a wider PHM programme can be summarised in the following flow diagram:
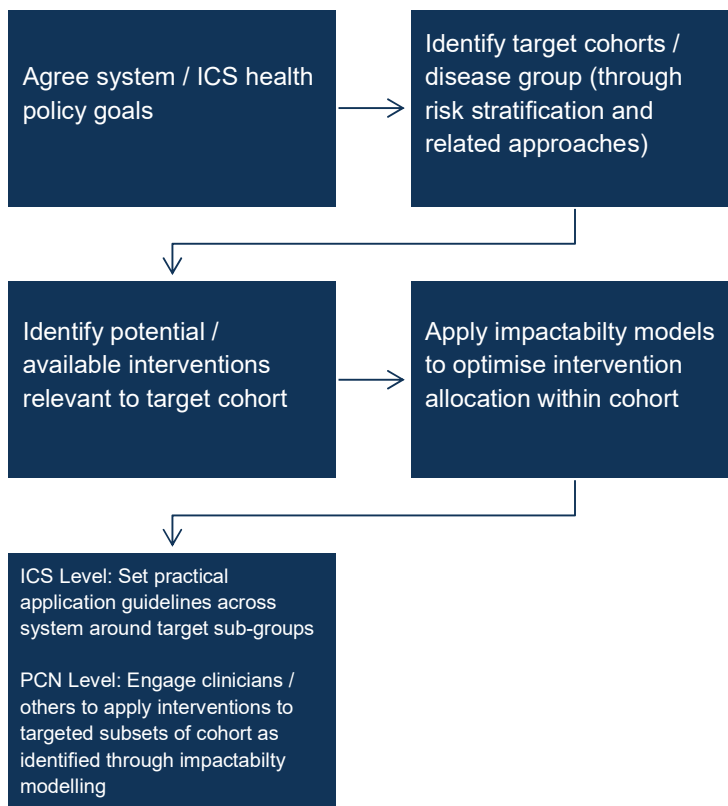


```
┌─────────────────────────┐        ┌─────────────────────────┐
│ Agree system / ICS      │        │ Identify target cohorts /│
│ health policy goals     │───────▶│ disease group (through  │
│                         │        │ risk stratification and │
│                         │        │ related approaches)     │
└─────────────────────────┘        └─────────────────────────┘

┌─────────────────────────┐        ┌─────────────────────────┐
│ Identify potential /    │        │ Apply impactabilty      │
│ available interventions │───────▶│ models to optimise      │
│ relevant to target      │        │ intervention allocation │
│ cohort                  │        │ within cohort           │
└─────────────────────────┘        └─────────────────────────┘

┌─────────────────────────────────────────────┐
│ ICS Level: Set practical application        │
│ guidelines across system around target      │
│ sub-groups                                  │
│                                             │
│ PCN Level: Engage clinicians / others to    │
│ apply interventions to targeted subsets of  │
│ cohort as identified through impactabilty   │
│ modelling                                   │
└─────────────────────────────────────────────┘
```

Figure 1: The role of impactability modelling within a PHM programme

Section 4. Practical challenges and considerations

Lead authors of Section 4: A. Shah, M. Flint, C. Martin, L. Morgan, C. Bull.

4.1 Research Methodology

The Working Party looked at the practical aspects of implementing PHM policies and interventions based on impactability modelling. We aimed to discover:

- How impactability models are currently used in practice by health care teams;
- What kinds of challenges with health care delivery we are trying to address with impactability models;
- How these models are developed and, in particular, what practical challenges exist with regards to how impactability is measured;
- What difference impactability modelling makes to the outcomes of preventative programmes;
- The ability and capacity of relevant health care professionals to build and use the models (i.e. technical skills); and
- What issues there may be in the availability and/or quality of data.

The purpose of the research in Section 4 was to identify key themes from current and potential end users of impactability modelling and document these insights as well as related advice and commentary.

In contrast to risk stratification tools, which appear to be better known (although not always fully utilised), knowledge and use of impactability modelling is comparatively sparse. This limited the research. However, the gap between theory and application also presents an opportunity for such modelling processes to be further developed and embedded within the UK health system's analytical toolbox. An approach to embedding these new tools is suggested by the wider change management process illustration in Figure 4 below.

Semi-structured Interviews

The first phase of the work involved semi-structured interviews with key personnel involved in the development, application or study of PHM programmes involving risk stratification and impactability modelling.

We used a mix of convenience and snowball sampling[3] to identify ten interviewees covering the characteristics in Table 3.

Interviews were conducted over the telephone between two interviewers from the Working Party and the interviewees, using a semi-structured interview questionnaire selected according to the background of the interviewee (Appendix 1). Two of the ten interviewees were interviewed together as they had worked on the same initiative.

A grounded thematic analysis of the interview notes was conducted, and the results circulated to the interviewees for further comment or revision.

---

[3] Snowball sampling is where research participants recruit other participants for a test or study (Statistics How To, 2014). It is used where potential participants are hard to find.

Organisation		
Academia	NHS	Other healthcare organisation
2	7	1

Expertise / Experience					
Academia	Data analytics	Clinician	Actuary	Healthcare management	Other healthcare organisation
2	3	1	1	6	4

Location	
UK	US
8	2

Table 3: Characteristics of interviewees

A literature review was conducted by searching the Embase database (a versatile, multipurpose and up-to-date biomedical research database) for "'impactability' or 'impactibility'" and then citation-chasing using reference lists and references supplied by Working Party members or interviewees. After summarising the included papers, the results were cross-referenced with the themes and concepts generated from the interviews.

During the interview process particular themes around current practice emerged. These were later synthesised and grouped under the following topics. We explore each in turn:
- Impactability modelling and risk stratification;
- Data issues;
- Analytical capability;
- Organisational issues; and
- Evaluation.

Recommendations that arose during the interviews are included in this section.

4.2 Practical Issues Raised in the Interviews
Context
Whilst risk stratification is widely used in the NHS, there is less application of impactability although there is increasing interest in the topic especially with the focus on the increased use of data to drive population health management.

Lack of awareness/understanding of impactability modelling
In practice people are often not clear on the definition of impactability, or its place in the assessment of an intervention.

Reflecting the newness of impactability modelling, there was not a consistent level of awareness and it appears that the use of impactability modelling is currently relatively ad hoc with pockets of good practice. The lack of understanding was illustrated by an anecdotal story whereby software providers were promising that their software would do "impactability modelling" but with no evidence to support such claims.

Short-termism

There is a focus for most NHS (and related) organisations on short-term results. However, improvements in risk stratification and impactability tend to accrue over a longer period. Even more so when conducting investigations or pilot studies with uncertain immediate results. Short-termism may also affect data consistency over time, which could in turn affect future data-driven processes.

Short-termism seems hard to resolve other than by convincing decision makers within the NHS to give greater consideration to the longer term, perhaps through the use of incentives

Timing of intervention

There is a generally held view that impactability may be influenced by timing, though the evidence for this is limited. For example, an intervention may be more likely to lead to behaviour change at the time of hospital discharge or at other 'impactable moments'.

Intangible disease sub-groups

Some categorisations of people are more concrete than others. Categorisation by disease is relatively concrete, but categorisations by severity are partially subjective. This affects impactability modelling, where there may be a need to identify relatively intangible categorisations like frail older people rather than people with diabetes who have had an amputation. People are affected by many different physical and mental conditions, and social factors, which interact with each other. Whilst in general it is easy to identify groups of people with a particular disease, identifying cohorts of people who are appropriate for a particular study, and the corresponding people to whom the results can be applied is much more difficult, as the data required is greater and may be less complete, consistent and precise.

Models developed for one sub-population may not work for another

Risk models that are fitted to data from a particular place at a particular time, or in a particular sub-group of people may perform less well when transferred to a setting with different populations, or at a different time when prior probabilities may have shifted, or the demographics may be different. This problem of a lack of transferability in geography or time can limit the re-use of risk-stratification, and impactability tools. A solution would be to hold multidisciplinary workshops (including clinicians, actuaries and other analysts) to ensure that the models continue to contain all the clinically significant parameters and are also calibrated with the correct underlying assumptions.

Anxiety

Risk stratification can generate anxiety in patients. The act of identifying someone as more vulnerable or at risk is naturally anxiety-provoking. This may have the advantage of acting as a driver for necessary change but can be a distressing outcome in itself, which needs to be taken into consideration. Anxiety is, in fact, often measured as a patient reported outcome of interventions. For example, it is particularly abundant around steps in breast screening and cervical screening. For professionals, there may be a fear that identification of gaps in care, a common part of impactability modelling, may be used to criticise them or might harm incentive payments.

Ethical difficulties

Some groups are difficult to impact, for example homeless people or those with severe mental health problems. Typically, screening programs send invitations to candidates by posting an invitation to an address, so homeless people are often tacitly excluded from the programme. Those with severe mental health problems are at higher risk of non-attendance and those who do not speak English may not be as able to respond to invitations. Unless care is taken to ensure otherwise, poorly constructed and implemented impactability models might exclude these categories as hard-to-reach, and therefore unlikely to benefit. However, it raises questions about abandonment of those with the most need, and the potential to exacerbate inequalities in care (Lewis, 2010). These points are further addressed in Section 6.2.

Patient Compliance

Compliance (or adherence) may be seen as a factor that might drive inclusion or exclusion from interventions. On the one hand some consider a high risk of non-compliance as a factor that reduces the likelihood of success of an intervention, but others might regard it as a marker of need that may itself be modifiable by an intervention (Lewis, 2010).

Data issues

Data is vital for the fitting of risk stratification and impactability models, and data issues featured prominently in the responses of all the participants. Many of the issues are not specific to impactability modelling alone, but remain important considerations nonetheless. Figure 2 summarises the data issues identified by the interviewees.
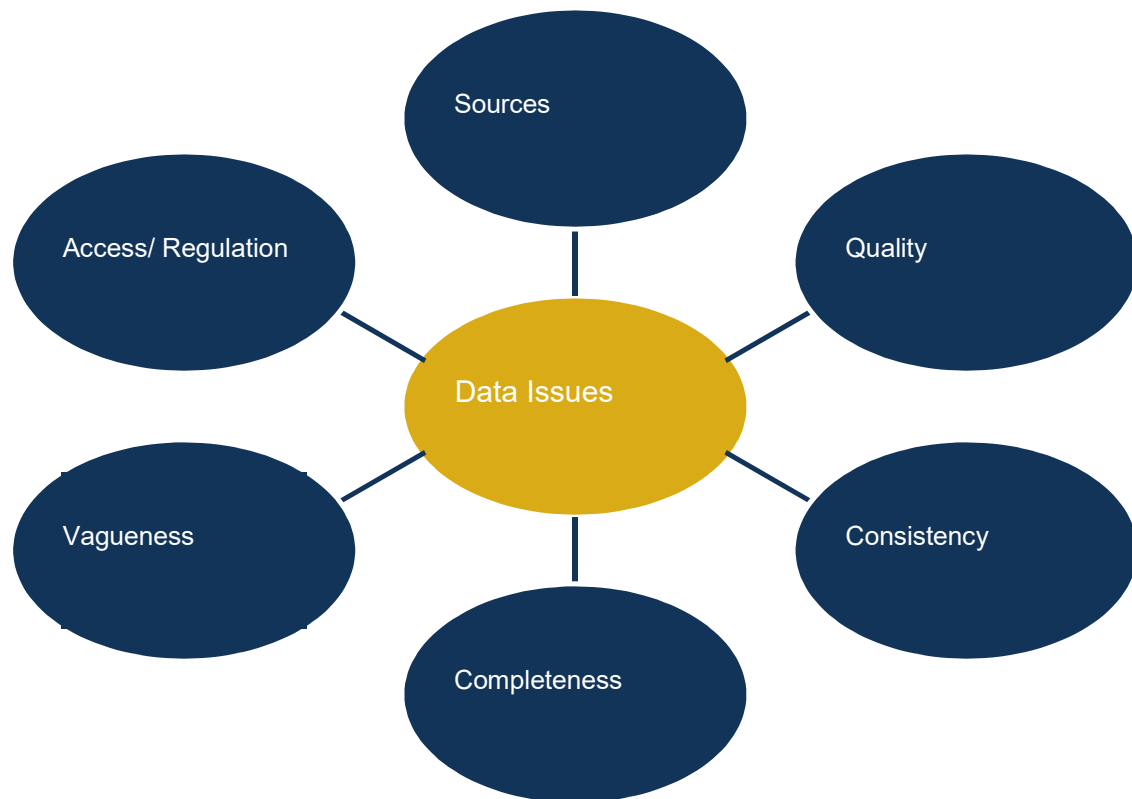

Figure 2: Potential data issues in impactability modelling

In practice the ideal data is rarely available however it would be useful to understand the limitations of different levels of data completeness for the purposes of PHM and impactability modelling. This should include data from social and mental health care services data in addition to primary and secondary care. Given the importance of social factors in impactability modelling, other useful sources of data might come from the Department for Work and Pensions for employment status or receipt of benefits, education to indicate educational achievement, criminal justice for a history of offending,

pharmacies who could confirm whether prescribed medicines have actually been collected, and opticians to indicate those who have poor vision. However, there are considerable ethical and data protection hurdles to be overcome (see Access section below). In practice there is a near unlimited range of data that could be used to build a more and more detailed view of populations and individuals (for example, data collected from wearable devices).

Data Quality
It is difficult to create a dataset that is sufficiently consistent, complete, unambiguous, and voluminous to be used for statistical analysis. This issue is magnified when considering across geographic regions or longer timeframes.

Consistency
There are various electronic healthcare records systems and standardisation is an issue when collecting data both within and across different systems. Different electronic healthcare records systems may structure their data in different ways. Different health care organisations may have different pathways for recording data, and different policies on how to do it. Different healthcare professions may have differing conceptual models of health problems and use different language to describe it. There are clinical coding systems to help mitigate problems like this such as the Read codes, which have now been replaced by SNOMED CT (a comprehensive, multilingual, clinical healthcare terminology resource). However, there are inevitably synonymous ways of describing a problem even within such coding systems, and these can be difficult to resolve.

Box 1: Data Sources

As the development of electronic patient records systems has matured, the number of databases of clinical and linked information has grown. Examples are given below:

1. The Clinical Practice Research Datalink (CPRD) database began in the 1980s with the aggregation of data from the VAMP medical GP records system. Since then it has been acquired by the Medicines and Healthcare Products Regulatory Agency (MHRA) and the scope of donating practice systems increased. Records are linked to other data such as deaths and Hospital Episode Statistics. 'The Health Improvement Network' database began as a branch from the original VAMP database and has similar, but less extensive content. It is now maintained by IQVIA.
2. Another database based around general practice systems is QResearch. This was developed using data from another vendor 'EMIS', and the data is currently under the care of a team at Oxford University.
3. UK Biobank recruited half a million people aged between 40 and 69 years between 2006 and 2010 (UK Biobank Limited, 2010). A wide variety of data, including genetic data was collected, with the subjects having their health tracked. It is a registered charity and the data is available to bona-fide researchers in healthcare and industry.
4. Activity and cost data at a local level is abundant, but there is often a lack of linked clinical or social care data.
5. The Secondary Use Service (SUS) is a data warehouse of patient-level information aggregated from a number of NHS funded health care organisations intended to support analytics in the NHS such as for planning and commissioning services.

Completeness

Social care data is often incomplete, and yet social factors are often very potent predictors of need. Consequently, a risk-stratification tool fitted to data may not perform as well as simple heuristics applied by clinicians who can observe relevant social factors such as living alone, lack of a social support network or poverty. This drives a subjective view that clinicians have a better understanding of social circumstances and may perform better than risk-stratification tools in some circumstances.

One practical solution to this is to apply statistical and heuristic approaches in a two-step process. The challenge then is to tailor the modelling to optimise the statistical results given the data available on application, and then to make the subjective decision making as well-informed as possible. In practice this would be very hard to optimise but the approach may allow positive use of partial data.

Completion rates are very low, and in some cases, are inversely related to need. For example, Medicaid enrolment forms in the USA have completion rates that are lowest for the most deprived people who have the greatest need. Variable incompleteness creates distortions and biases in the data. One, widely used, algorithm for selecting patients with complex health needs in the USA under-selects black people with complex health needs compared to white people, as the algorithm was based on costs rather than direct measures of health.

Service users may access services across different providers, and so their data may be spread out across different electronic healthcare records systems if it is not shared effectively.

Sometimes there is no data. If a particular risk-stratification or impactability model requires a particular field of data, it is unlikely that field will be populated in the records of all people.

There is sometimes an expectation that data would be collected for the purposes of subsequent analysis. However, it needs to be borne in mind that the primary purpose of an electronic patient record system is to support the continuing delivery of care to each patient. This means that data collection for analytic purposes is of secondary importance in what are often very time-pressured interactions. Even when a particular data field is required for clinical care alone, it is recorded from the perspective of the delivery of care rather than support for analysis. This can subtly affect the structure of the information rendering it less reliable for research.

Vagueness

Some things are hard to measure, or only exist in intangible forms like degrees of severity, such as social status or theoretical subgroups of disease groups. For example, a 'brittle diabetic' is a person with diabetes with high variability in measures of control that can rapidly change from good control to instability very quickly, and 'resistant hypertension' indicates a hypertensive with a poor response to several different forms of treatment. The meaning is dependent on what the algorithms and thresholds would be for deciding if a particular person with diabetes was 'brittle' or a particular person with hypertension was 'resistant'. The vagueness does not only apply to diagnosis. Some interventions may seek to identify lonely people, or those who are 'frail', both of which have significant degrees of subjectivity.

In the analysis, we distinguish between vagueness and consistency. Consistency is about lack of variability; vagueness, about imprecision. For example, we can measure systolic blood pressure very precisely. It is not vague, but if we measure it three times in quick succession, it will almost always give different results. So systolic blood pressure is inconsistent but not vague. However, if we do an x-ray to assess the severity of a broken bone, we will only get an imprecise or vague measure of how bad it is ("spiral", "comminuted", "greenstick"), or a mix of features to different degrees, but if we repeat the x-ray, it will still look pretty much the same. The x-ray of a broken bone will be consistent, but vague. When developing impactability models, the consistency and vagueness of the data will need to be taken into account.

Access
Types of data
As mentioned in the interviews, the most consistently available data for PHM is activity data. Contacts between service users and NHS organisations are routinely collected, as are admissions and procedures. This data is certainly valuable for PHM, but there is a need for clinical and social care data also. Figure 3 summarises the different data mentioned for impactability modelling.
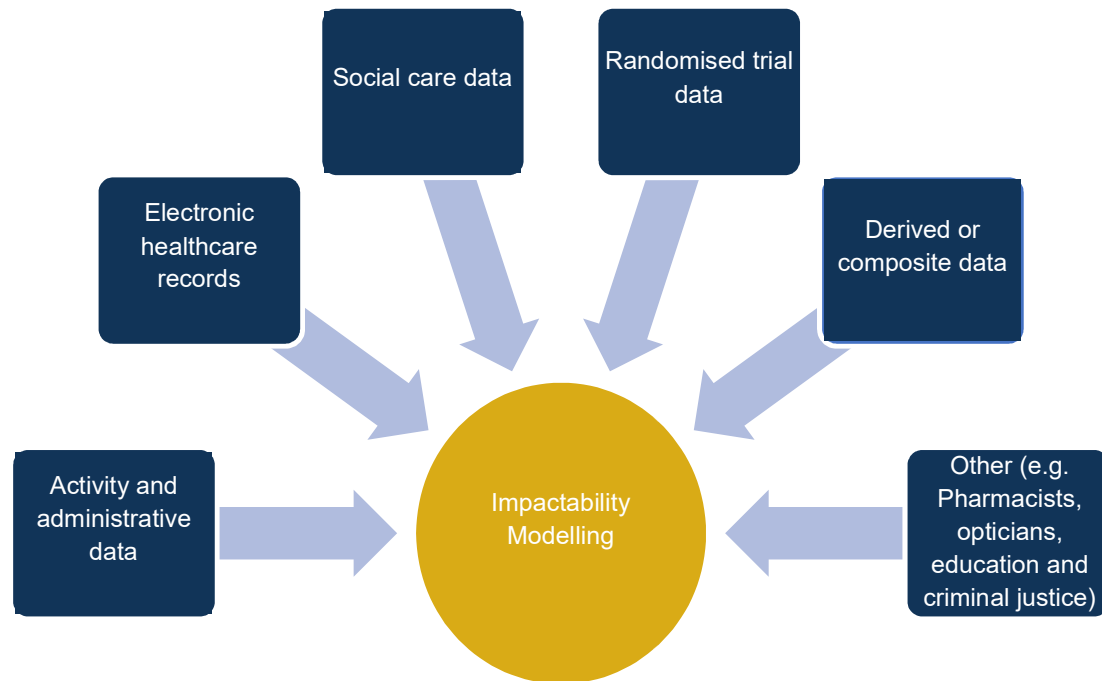

Figure 3: Potential types of data used in impactability modelling

Activity and administrative data
This data consists of records of contacts with healthcare, service utilisation, operations and other treatment episodes. This is often available at an aggregate level for an organisation or a district but is potentially available as linked data at the individual level. The aggregate data is relatively non-sensitive and is often routinely available in the UK. Individual data is hard to access, even when anonymised.

Electronic healthcare records
This is data from the individual patient clinical record. The minimum data sets would normally consist of diagnoses and prescriptions, but may also include biometrics like BMI, vaccinations, smoking status and other lifestyle factors, treatment episodes, and narrative descriptions of their care. This data is potentially extremely sensitive and is therefore difficult to access.

Randomised trial data
Data from good-quality randomised trials would be the ideal kind of data for impactability modelling as there is a valid comparator group with control of potentially confounding factors. However, this kind of data is rarely available (Lewis, 2010, Lewis, 2015).

Social care data
Several interviewees identified socio-economic and social care data as being potentially the most valuable for impactability modelling and the majority of the most impactable patients in one study had at least one significant social determinant (DuBard and Jackson, 2018). However, such data was considered to have very limited availability because of a lack of interoperability between systems, limitations in recording or restrictions on access because of data protection concerns.

Derived or composite data

Some data used in impactability modelling consists of variables derived from other data sources. These may be risk scores such as QAdmission that quantify the risk of admission based on a variety of demographic and clinical data (Hippisley-Cox and Coupland, 2013), clinical population stratification or clusterings such as the Adjusted Clinical Groups (ACGs) from John Hopkins University (Buja et al., 2019, John Hopkins ACG, 2018). The ACGs are widely used in the USA for population profiling, predicting costs, performance analysis and case management.

A similar approach has been taken in marketing for many years, and there are population segmentation products designed for business, but which may have application in PHM like Acorn or Mosaic (NHS England et al., 2018, Vuik, 2017).

Other sources

Other sources of data were suggested by interviewees as being potentially predictive of the success or failure of intervention or for shaping the type of intervention for individuals, such as data on the collection of prescriptions from pharmacists, educational achievement, and criminal convictions, though the ethical and data protection issues might be insurmountable. These data might be markers of the ability to adhere to health interventions or messages, or of a history of substance abuse or chaotic lifestyles (DuBard and Jackson, 2018, Freund et al., 2012, Prusinski, 2017).

Regulation

Access to data appears to be poor, even when it exists and is theoretically available. Access to individual clinical data hinges on the consent of those donating data (e.g. BIOBANK, CPRD etc), and sometimes the consent will cover some purposes such as medical research, but not others such as non-medical research. Mistrust between controllers of data and the users of data, particularly over the motives for access, can be a barrier.

Even when the purposes for accessing the data are legitimate, the fear of falling foul of data protection regulations or media hype can cause data controllers to err on the side of caution and over-restrict access. This may arise from concerns about public opinion, regardless of the legitimacy of use, or from a lack of detailed understanding of what the regulations do and do not permit.

Clinical records contain very sensitive information about individuals and it is important that health services can be trusted by patients to keep that information confidential to ensure disclosure of information necessary for healthcare. This may include research for the purposes of PHM, including risk-stratification and impactability modelling, which would be viewed as a legitimate use given ethical research committee approvals. Any application for access to data needs a well-crafted business, ethical and research case. Analysts need to make it clear that their purpose is health services research only.

The anxieties of data controllers may extend beyond regulatory requirements and media hype. Much of the analysis that goes on in health care is for the preparation of statistics for performance management, and data guardians may need reassurance that the data accessed will not be used in this way.

Difficulty searching information

The data structure of clinical records systems is necessarily extremely complex. This contributes to the difficulties encountered in inter-operability, but also presents challenges in searching for and extracting data in a reliable fashion.

Additional recommendations on data

Further considerations for improving and standardising access to data should include:

- Prepare a well-crafted business case and engage with stakeholders and ethical reviewers early for accessing data to gain the trust of data guardians;
- Build an open and transparent compliance infrastructure for data handling; and
- A phased approach to systems integration across organisations may have a greater chance of success than a waterfall approach where all systems change at once.

Analytical capability

Context

To carry out effective impactability modelling the right analytical skills are needed. There is huge demand for analytic services in healthcare, but a variable supply of skilled analysts.

Current analysis is carried out by a mix of internal analysts across the NHS, academics and commercial consultants.

Availability of resources

In general, the interviewees we spoke to did not suggest that it was necessarily the availability of analytical skills which was preventing impactability modelling but rather issues around awareness, organisational priorities and access to appropriate data. However, some believed that analytical time was being diverted, for example, towards the preparation of statutory returns rather than generating output designed to improve the health of the population. This suggests that some NHS bodies had limited appetite to fund impactability modelling where there may not be an immediate benefit.

Analysts are not the only important resource in designing impactability models. There are arguments that analysis should be a collaboration between analysts, clinicians, health care managers, ethicists and patients. Clinicians and patients may have useful insights into the most effective data elements to include in a model (Steventon and Billings, 2017).

Handling results

Even when analytical skills are available, there may be difficulties in communicating results. Those who are skilled at analysis are not necessarily skilled at communicating the conclusions of it, and this is a vital step in the process of maximally exploiting the power of data and information. Data visualisation capacity was highlighted as a particular need.

Common pitfalls in analysis

Analysis or modelling of a programme or intervention might focus on immediate processes and outcomes of interest. However, it is important to properly understand the intervention and consider down-stream effects by considering changes to the system over time. In practice downstream effects are very hard to quantify in a manner that makes them comparable to more immediate effects.

New methods such as artificial intelligence (AI)

There is an expectation that new approaches like artificial intelligence will be able to perform better than traditional risk stratification models, and it is mentioned in the NHS Long Term Plan as a practical priority of the NHS digital transformation (NHS England, 2019). However, a comparative analysis of regression, artificial neural networks, decision trees and cluster-analysis found that regression methods probably performed the best overall (Vuik, 2017). The application of AI is only likely to yield results in the longer term. Also, many of the problems highlighted here will be equally damaging to AI modelling as they are to traditional modelling methods. Resolving current issues with traditional modelling will help solve future AI modelling challenges. For more on implementation considerations, see section 6.2. Ethical concerns may be amplified by data-hungry AI.

Additional recommendations for analytics

- If impactability modelling is to become mainstream then analytical teams will need to be scaled up. Although some modelling applications may suit analysis on a national scale it is likely to require more local involvement across the UK at least in part because some data is only held on local systems. The development or enhancement of 'hubs of excellence' would allow skills and resources to provide analytic services to organisations across the NHS.
- As well as increasing the number of analysts working in this area it would also require the broadening of analysts' skillset. In particular this would be needed to communicate the story behind impactability modelling and the benefits it could bring. This might require stronger report-writing or data visualisation skills and verbal presentations.
- Appropriate training should be provided to the end-users of risk-stratification and impactability modelling tools to ensure, engagement and appropriate application.

- Encourage greater collaboration between analysts and clinicians to develop and refine impactability models and implement with patient care. I.e. a cultural change to get clinicians and analysts working together more closely together with some upskilling at the clinical 'coal face' so the outputs from these tools is appropriately considered at the point of intervention or 'impactable moment'.

Organisational issues
Context
The NHS Five Year Forward View of 2014 and its update in 2016 increased the emphasis on prevention and integrated working between different parts of the NHS, and with social care and local authorities. A series of fifty 'Vanguard' projects were set up to trial innovative ways of working. This was followed by the development of 'Sustainability and Transformation Partnerships' ("STPs") between the NHS and local authorities to help translate the local successes of the Vanguards to general implementation across the country. Some areas formed 'Integrated Care Systems' ("ICSs") with even closer collaboration which included voluntary, community and social enterprise sectors, and with greater control and responsibility over budgets. The NHS Long Term Plan set out an aim that every part of England will be covered by an ICS by 2021. PHM initiatives targeting those with most to gain from interventions delivered by integrated multi-disciplinary teams, with a particular emphasis on prevention, is a central part of these reforms.

STP's in formulating their response to the NHS Long Term Plan should make specific reference to PHM. Indeed, PHM, including risk stratification and impactability analysis, are core pillars necessary to realise the benefits that great collaborative delivery of healthcare which STPs and ICSs will enable. One aspect of the increased collaboration across organisational boundaries within a STP area will be the ability to plan interventions which can seek to realise benefits across the health system, not just for individual organisations who deliver them. Impactability modelling is essential to understand and optimise these benefits and STPs provide the ideal organisational framework to deliver these changes effectively.

Practical Issues Raised in the Interviews
International context
Whilst the focus of this exercise is the UK and the NHS, there are lessons from other countries and other models of health care delivery. The organisational or funding structure may affect impactability. For example, in insurance-based systems, the access to an intervention may be complicated by issues of insurance status.

Relationships
Morale and the relationships within and between teams in the NHS, social care and other organisations is critical to success. Multidisciplinary teams are hard to implement when relationships between the providers involved are poor. The success of an initiative requires that those implementing it have the trust and motivation to engage with the process. Otherwise NHS organisations do what they are directed to do but may be mechanical and not necessarily creative in solving problems. Some interviewees reported poor relationships between different providers and how this constrained the ability to effectively collaborate across a system. However, it was also recognised that some organisations may be stretched to meet current demand with limited time or resources to allow individuals to explore greater collaboration.

Trust is an important issue and there is a general feeling that locally developed initiatives are preferable to top-down, centrally promoted ones. Those that are implementing programmes locally or providing data must have the trust and motivation to engage.   Some of the system-wide benefits of an effective PHM strategy will require interventions at one point in the system to realise benefits elsewhere (for example, earlier social care interventions which may lead to less deterioration in health conditions and lower secondary care attendance). Trust across provider organisations within a system both at the organisational level and at the individual level between clinicians and other

professionals is essential to enable these cross organisational interventions to be identified, implemented and the resulting benefits realised.

What PHM means to the clinician at the point of care

PHM is being directed for implementation centrally, but there may be only a rudimentary understanding of what that means to those working locally. In reality PHM implementation has sometimes been limited to risk stratification.

> **"In my limited experience, I observe a tendency to present population health interventions in binary terms - either an intervention is data driven and the attendant health expert is just following rules or it is heuristic"** - Interviewee

Ideally interventions would be optimally targeted by using the results of data analysis to inform the subjective decision making of the attendant health expert. This binary view is presumably not helpful in getting the buy-in of practitioners. Section 6.2 provides further discussion of the fairness of an intervention, which depends on how interventions are applied, potentially using a sliding scale across groups rather than making a binary decision.

It was also recognised that there is a balance to be struck between PHM and impactability modelling 'dictating' the response for individual patients and the role of clinical judgement at the point of care. Effective PHM implementation requires both a greater understanding by clinicians of the approaches and benefits of the analysis performed and ways to get the resulting insight into the hands of clinicians at the point of care to inform the heuristic decision-making approach. The exact response will depend on local circumstances, clinical buy-in to the broader PHM approach and a range of other organisational and practical factors.

Inter-operability of records systems

Inter-operability of information systems is a recurrent theme. NHS organisations are often large and complex. Additionally, linking NHS data with data from non-NHS organisations, such as Local Authorities, is even more challenging. Clinical information systems will also be large and complex and will continuously form a central part of interactions with patients. An NHS body will orientate itself around one records system which then becomes an embedded part of the organisation. This structural orientation around a system creates a barrier to change and may affect the ability of different organisations to work together.

Some of the benefits of impactability modelling across a population require an understanding of how individuals within that population interact with different parts of the health system. This requires the connecting-up of data about those individuals across system providers.

Patients will often access care at different locations with different records systems such as their GP, the local hospital, a specialist hospital, the pharmacist, dentist or optician. Whilst some systems attempt to span these different contexts of use, there has been limited success so far in implementing systems across all suppliers in an area, particularly between primary and secondary care.

Organisational factors influence impactability

Impactability modelling should preferably include characteristics of the organisational structure as well as characteristics of the patient. The way an organisation is structured necessarily influences its ability to co-ordinate care in particular ways (Sheaff et al., 2015). For example, organisations not orientated around primary care may have problems identifying and reaching target patients. If an initiative requires delivery of care in a certain format, the means with which to deliver that care needs to exist wherever the initiative is to be implemented. This might reflect variation in structure, or variations in health priorities in different areas. In healthcare systems outside of the UK, it might relate to the supplier of healthcare that is accessible by the patient, or their health insurance status (Prusinski, 2017).

Performance measures are also an issue. Where a health intervention provides an overall benefit but adversely impacts the performance measures of one provider it can be very hard to obtain the buy-in of that provider.

Training and understanding

Introducing a new healthcare intervention or set of interventions inevitably has training needs associated with it. This includes ensuring those needed to successfully implement a programme of care understand how their participation and contributions impact population health more widely than the context of care within which they work. PHM, risk stratification and impactability modelling are not necessarily readily understood by people working at the front-line of healthcare. Ensuring they understand the motivations for the work, and how it leads to better outcomes is important to securing their full co-operation. This also extends to providing visibility of the cross-system benefits which an intervention can have but which may not be immediately visible to those who are tasked with delivering the intervention.

Continual reform

A tacit theme that has emerged through the interviews is that of the difficulty of building new ways of working in a constantly changing environment. This has emerged from the consistent backdrop of implementing change and organisational factors affecting data accessibility. Major reforms in health and social care is now almost a continual process (Nuffield Trust, 2019). Priorities and initiatives change, and sometimes the data may change with it. In order to preserve the incremental improvements achieved with each creation or revision of a service, the elements of that service need to have longevity independent of the organisational context within which it is applied. Risk-stratification tools or impactability models that are embedded in the idiosyncrasies of a particular organisational context may not survive transfer to another organisation, or another time. In the current context, the NHS is going through a significant period of organisational change. The development and implementation of impactability models to drive the delivery of interventions to those who most need them will need to endure through these changes such that the benefits are realised. This is especially so for benefits expected to be realised in the longer term. Measuring the benefits of interventions is an important part of assessing the success of their impact. This requires a baseline against a stable pathway of care from which to measure the benefits of particular changes in care or other interventions. Continual reform limits the ability to establish a baseline and isolate the impact of an intervention from other system changes.

Embedding impactability in a wider change management process

To enhance the ability to make data-driven decisions, Impactability modelling could be used more systematically and become an important tool in the NHS's management framework. The adoption of such tools/modelling in a systematic way would require a change management process to incorporate the modelling, as well as implementing the results i.e. implementing the optimal set of healthcare interventions as suggested by the modelling results. Change management is an ongoing process: ultimately, ongoing monitoring and evaluation of both the models and the interventions would allow for a continuously improving feedback. Figure 4 illustrates the steps and nested subsets for managing the introduction of impactability modelling and its results (adapted concepts from (Dalal et al., 2018).
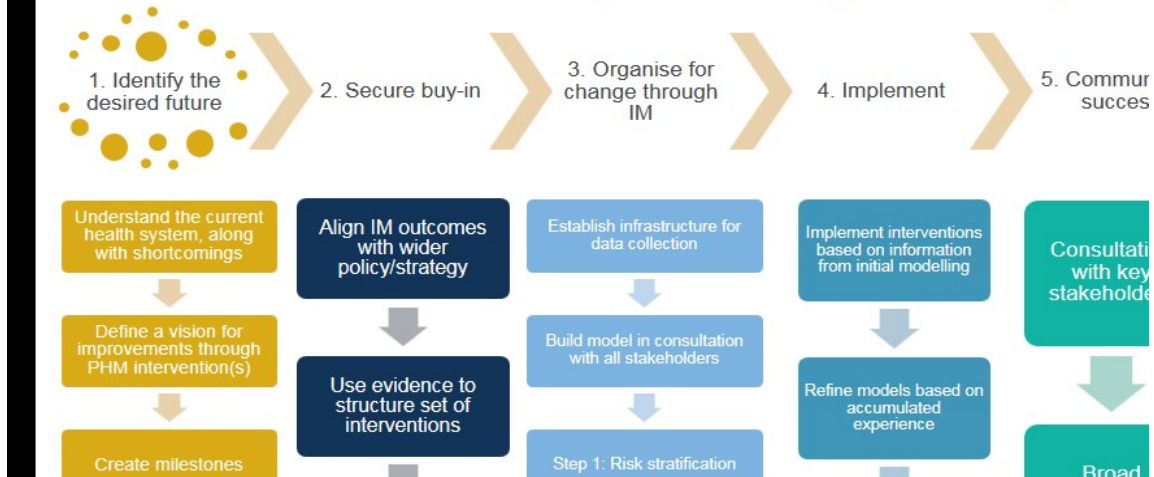
Figure 4: Impactability modelling as part of a continuous change management process

Additional Recommendation

To help build trust and co-operation between stakeholders we would recommend:

- Providing support for organisations in understanding segmentation, risk stratification and impactability modelling. This may increase trust and access to data and improve co-operation between individuals and organisations.
- Where successful case studies exist, these should be widely published so that others can learn.

Evaluation
Context

Evaluation of modelling is important to ensure that its introduction is having the intended effect, and is not giving rise to any unintended consequences such as paradoxical increases in costs, nock-on decreases in health or quality of care, or exacerbations in inequalities. However, unlike traditional medical interventions there are no standard evaluation procedures e.g. randomised controlled trials. In adopting new modelling, it will be necessary to show that the intervention works - and ideally that it continues to work as the system evolves.

The evaluation of a proposed intervention should include careful health-economic modelling and evaluation with 'Real-World' evidence to ensure the expectations of benefit in cost or improved health are realistic, and the costs of the programme can be weighed against the opportunity costs of alternative investments in care. Some interventions may be subject to NICE guidance where a careful, structured and evidence-based evaluation including economic-modelling is required.

Practical issues raised in the interviews
Holistic approach

Ideally the evaluation would encompass all effects arising from an intervention and assess the benefits and costs in a consistent and robust framework.

A thorough understanding of the structure of an initiative, and the care pathways it entails is therefore needed for a proper evaluation. Focussing on specific parts of the process in isolation may not capture important collateral or paradoxical effects. Where data mining approaches are used (with models based predominantly on data rather than theoretical considerations) then extra care will be needed that true benefits will be achieved.

Outcomes

It is important to keep in mind the overall goals of the project when selecting outcomes for evaluation. Many things may be easy to measure but may not be closely associated with the intended outputs of the project. For example, adherence to treatment is frequently measured as an outcome rather than the objective of the process. A failure to comply does not necessarily indicate continuing risk but may indicate recovery. However, some outcomes cannot be realistically measured, and it may be necessary to resort to proxy measures. For example, a programme that aims to reduce the risk of heart disease or stroke in diabetics would need many years to accumulate the outcome data on strokes and heart attacks for analysis, in which case it may be necessary to rely on proxy measures like reductions in blood pressure, cholesterol and smoking.

It is important to consider the perspective of the analysis in selecting performance metrics of risk-stratification tools and impactability models. At a population level we might be concerned with the proportion of all cases identified (sensitivity) or the true negative rate (specificity) of a tool, but from the individual perspective the probability that a particular positive or negative result is correct is more important (positive and negative predictive value).

> *Box 2: Example of the importance of measuring "real-world" outcomes*
>
> PRISM (Predictive Risk Stratification Model) is an Emergency Admissions Risk Prediction (EARP) model developed in Wales and commissioned by the Welsh Assembly. It stratifies the population into four groups according to the risk of emergency admission. Its discriminative performance was good with a c-statistic (area under the receiver-operating-characteristic curve (ROC)) of 0.749. However, the results of a cluster-randomised trial of its implementation in South Wales had the opposite of its intended effect. Emergency admissions, emergency department attendances, GP event days, outpatient visits and overall costs all increased, whilst satisfaction reduced (Snooks et al., 2018). This may relate to the uncovering of unmet need. The application of impactability modelling may have enabled the programme to be better targeted at those who would truly benefit, and might have had a different outcome.

Pilots

Piloting of proposed initiatives is necessary. It is difficult to know how risk-stratification, impactability modelling and any interventions will perform until applied in the real-world, and how a programme operates in the context of a large and complex organisational structure with competing demands on its time and attention. Lessons learned from piloting a proposed initiative can lead to revisions that can turn from one doomed to failure to one destined for success (Snooks et al., 2018).

Evaluation of programme should be sufficiently complex to capture all effects that might arise, including prior and linked events. Focussing on an isolated component, at a single point in time may fail to capture all consequences that are observed.

The most common approach to evaluation of a PHM initiative is the before-and-after study, where outcomes are measured before and after the implementation of a change. Whilst it should be

expected that a truly successful intervention gives rise to a measurable change, it is impossible to infer causation in this type of study, as there are many other uncontrolled factors that might have occurred to give rise to the same observations, and the statistical phenomenon of 'regression towards the mean' would suggest that a high risk group at the start of a time interval would, on average, experience a reduction in risk over time regardless of any intervention (Galton, 1886) See Box 3 below. The only reliable test of causation is the randomised controlled trial, which is often impossible or impractical to implement in this context (Lewis, 2010, Roland and Abel, 2012). Instead of a before and after trial it was suggested that evaluation against a paired sample was more reliable. However, there are often difficulties in identifying a similar enough geographical dataset to make comparisons against.

Whilst impactability is important it is only a component of the initiative as a whole, and more specific evaluation of the impactability component is needed including measures of sensitivity, specificity, positive and negative predictive value, and the performance in comparison to the main alternatives such as referral from practitioners or simple risk stratification.

Box 3: Regression to the mean

In the 19th Century, Francis Galton noted that the offspring of plants that had large seeds, tended to produce seeds that were smaller on average, and that the same was true of height in humans. He also noted that the change in the average value between generations was proportional to the deviation of the parents' measurement from the average. He initially mistook this as being a phenomenon related to inheritance before realising that there was a simple statistical explanation he referred to a 'regression towards mediocrity' (Galton, 1886).

'Regression towards the mean' is a statistical phenomenon describing the observation that if a measurement is taken of something inherently variable and found to be high or low in comparison to the average, then if that measurement is repeated, it is more likely than not to be closer to the average than the original measurement.

In a before and after evaluation of an intervention to reduce admission rates in a sample of people with high admission rates, regression towards the mean would lead to an expectation that the admission rate in that group of people would be lower regardless of the effectiveness of the intervention.
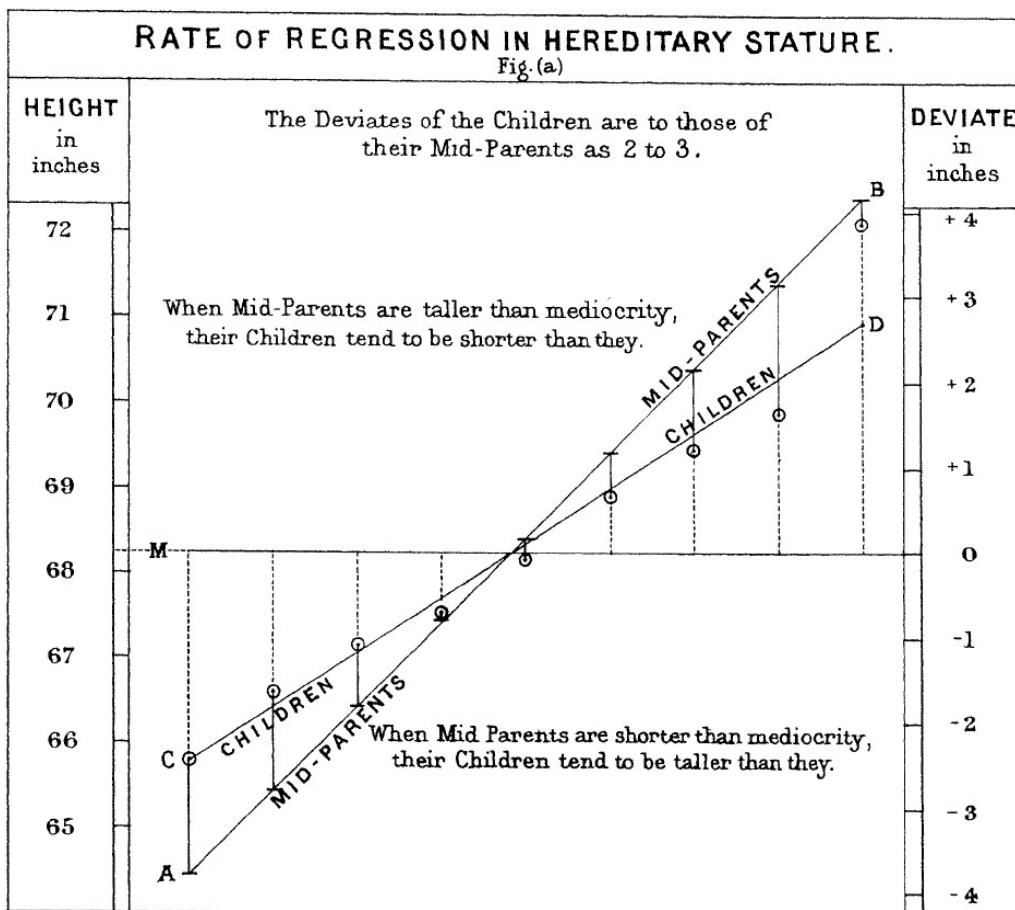

Figure 5: Rate of regression in hereditary stature taken from Galton 1886.

Additional recommendations

- When calibrating impactability models it is important for analysts to consider the objectives of the programme. The thresholds for action in a project aimed at reducing costs will differ from one aiming to maximise aggregate health status or cost-effectiveness. If an objective is to reduce inequalities, this may have implications for the mechanisms of case selection and impactability modelling.
- When applying impactability analysis, it is useful to also think about modelling 'impactable moments' i.e. if the intervention (or set of interventions) is applied at the optimum time, the impact can be much greater.
- Use of artificial intelligence or data-mining methods may give rise to improved performance in risk-stratification and impactability modelling.
- To alleviate the problem of changes in performance of risk-stratification tools over time and between different contexts, it may be possible to develop dynamically updating models that regularly refit to collected risk factor and outcome data over time. This would not only involve analysts, but would need processes embedded in the organisational workflows and IT systems to feed the data to the evaluation process.
- Consider how the evaluation of a new service may need to take account of how the application of impactability modelling may change the content of the service user population, and potentially changing the expected mean outcomes.

Section 5. Potential approaches and models

Lead authors of Section 5: D. Beddows, I. Bakbergenuly, M.R. Elsheemy, T. Hayward, J. Robertson, J. Seymour.

5.1 Introduction

As per the interviews in Section 4 , there are many different impactability models. Within each impactability model, there may be multiple statistical methods that could be applied. This section of the report considers both the type of impactability model and also the range of statistical and data science techniques that can form components to help develop those models.

To be clear, this report does not recommend the design of any particular healthcare intervention or attempt to judge the clinical effectiveness of any intervention. We are making the distinction between "impactability model" and "statistical method" for the purpose of highlighting that there may be a number of competing statistical methods depending upon the features of the impactability model.

Impactability model types

There are multiple ways of approaching impactability modelling, for example as outlined in Lewis (2010). We have decided to take the approach of identifying an initial list of models via the literature and interviews (key examples of which are listed in Table 4 below), and classifying each model into one of three categories: "traditional", "statistical" and "survey-based":

- The **traditional models** are ones which individual clinicians and health professionals already follow for their individual patients - they are based on reviewing information held in a patient's medical records. With improvements in electronic data recording and accessibility of this data, there is potential to harness the datasets to more quickly apply these approaches to segments or whole populations rather than individual patients one-by-one.

- The **statistical models** are ones which have been developed to apply advanced statistical techniques to patient data that is often readily available on Electronic Medical Records (EMRs) and hospital data systems. These are methods which require underlying mathematical or statistical models, calibrated for large datasets.

- The **survey-based models** are ones which require collection of additional data which is not typically captured on EMRs or hospital systems on a routine basis. This data is typically collected via patient questionnaires and this practically means that it is limited to smaller groups of patients than the other approaches. These approaches also involve advanced statistical techniques.

Examples of impactability models		
Traditional	**Statistical**	**Survey-based**
Prioritise patients with: • high gap score • high weighted gap score • one or more ambulatory care sensitive conditions De-prioritise patients with: • stable characteristics • extremely high risk Impactable moments • e.g. post discharge from hospital	Risk stratification / risk scores Rising risk score (rate of change of risk score) Comparing service utilisation, risk characteristics and condition severity against benchmarks Data science methods: clustering; decision trees; neural networks.	Questionnaires/frameworks for assessing individual patients, e.g.: • Patient Activation Measure (patient readiness to manage their own care) • ANGEL score (qualitative assessment of the complexity of patient need)

Table 4: Examples of impactability models by category (descriptions are provided in Appendix 2)

We have adopted these three categories in an attempt, at least in part, to reflect the nature, quality and quantity of data needed for each model. However, it should be noted that a certain amount of data will be required to derive any quantitative measure of impactability, under any of these categories. It is important to be aware that incomplete datasets could lead to unintended biases in the results; for example, restricting service utilisation data to hospital activity means the outputs are only based on the cohort of people who were ill enough to attend hospital during the data observation period.

Selecting a short-list of impactability models for deeper study
We narrowed down the initial list of models with a view to evaluate four impactability models, spanning all three categories. The chosen short-list included:

- Model 1: Prioritise patients with high gaps-in-care scores
- Model 2: Condition severity benchmarking
- Model 3: Propensity-to-benefit scores
- Model 4: Patient activation measures

For each impactability model there could be a number of different statistical methods. E.g. The gaps-in-care model and propensity-to-benefit model are 2 different types of impactability model whereas the choice of Generalised Linear Model used to derive the propensity-to-benefit model is the "statistical method".

Evaluation criteria for impactability methods
Discussions of the evaluation criteria provided the rationale for focusing on the chosen models. These evaluation criteria are detailed below. The criteria are grouped into themes, broadly aligned with the "quadruple aims of health care" (Sikka et al., 2015) and a further theme covering additional technical concerns.

Patient Outcomes & Experience	Healthcare Utilisation
Which outcomes are most interesting? Can they be measured? How can they be improved? What time frame should be considered? What are the measures of success? Does the impactability assessment result in patients needing more appointments / procedures / diagnostic tests? Impact on health inequalities	Resources needed for performing the impactability assessment and the resulting net effect on clinician/system resources of implementing the results. Burden of data collection & analysis Types of costs: staff involvement; clinicians training; Raising awareness. Is the necessary technology available? Is there sufficient processing power of hardware? Do people know how to use the technology? What would be the impact on health inequalities?
Healthcare Professionals Experience	**Technical Concerns**
Does the method lead to higher or lower workload for health professionals? Are staff sufficiently trained in the procedures / techniques / understanding results of the analysis? Is recruitment of new staff necessary?	Is the necessary data available (e.g. sufficient history, individual level)? How quickly can it be accessed? Data quality Data privacy, information governance, GDPR, etc. Do the statistical elements of the method have solid theoretical bases? Are they clinically sound? Are the methods transparent? Have they been peer reviewed? Are assumptions justified? Are issues such as bias and regression to the mean a factor? Can actuaries add value to the field regarding each model? E.g. the statistical approaches would tend to lend themselves more to actuarial skillset.

Table 5 Evaluation criteria for impactability methods

5.2 Types of impactability model

The following sub-sections consider each of the four short-listed impactability models. They give a description of each model, the advantages and disadvantages, the data requirements and other important considerations.

5.2.1 Impactability model 1: Prioritise patients with high gap score

NICE guidelines are evidence-based and set out the care and services that are suitable for most people with a specific condition or need. The objective of these guidelines is to "prevent ill health, promote and protect good health, improve the quality of care and services, promote integrated care where appropriate and help health and social care professionals adapt and provide health and social care services" (NICE, n.d.-a).

An impactability modelling approach for PHM.

Impactability modelling aims to identify gaps in treatment actually provided compared to that recommended by the relevant guidelines (such as NICE guidelines in England, or other measures of optimised care). Patients with the largest gaps in care can be prioritised and it is assumed that, by providing these patients with more of the recommended treatments, their health will be impacted. Note that health commissioners (such as Clinical Commissioning Groups in the NHS) may not always adopt standard guidance for commissioning services in situations in which this is justified.

For each condition and set of circumstances, the list of treatments recommended by NICE could be compared to the list of treatments a patient actually received over a specific period. If a particular recommended treatment was administered, the 'gap score' for that treatment would be 0 while if it was not administered, the gap score would be 1. The gap scores for all recommended treatments

could be aggregated for each patient. Patients with high gap scores would be identified and prioritised within the particular PHM programme being developed.

A statistical model is formulated in Appendix 3.

The model can be refined by applying weights (i.e. values other than 1 or 0) to each gap.

This impactability model is a PHM approach in the sense that it helps to allocate scarce resources to particular individuals in the population using evidence-based guidelines to inform the allocation decisions. It guides allocation of resources to individuals within the population. There will remain individuals with gaps in care but the model aims to ensure that those with the largest gaps are addressed first.

Advantages and disadvantages for model 1
Advantages
- This model could highlight particular factors that are contributing to high gap scores and inform systemic changes to reduce gap scores at a system level. For example, it could be that NICE guidelines are not being followed for particular types of patients (e.g. patients of certain ages) or it could be that they are not being applied within certain hospitals or regions.
- Potential to improve equity since it encourages consistency in prescribed treatment and prioritising patients with the highest capacity to benefit. Furthermore, where gaps in care are more common in areas of high deprivation, this approach would be expected to reduce health inequalities.
- Recommended treatment based on evidence-based, clinically robust NICE guidelines.
- Ability to measure effectiveness of implementation since the change in the 'care gap' would be directly measurable.

Disadvantages
- This model considers gaps in care but does not consider unwarranted care. For example, if patients are receiving treatment that is considered wasteful and not recommended by NICE, this will not be accounted for using the gap score modelling approach.
- This particular method does not consider patient outcomes since the focus is on the treatment that patients receive rather than how they respond to it.
  - Outcome measures of interest could be defined and measured at time t and time t+1 to determine if the programme has been a success.
- Relies on having a guideline available for a multitude of scenarios and defining optimal treatment for a large number of sub-groups. This may result in an overly complex model structure and lead to difficulty in interpreting the clinical validity of the results.
- If the guidelines do not consider all appropriate factors, the assigned 'gap in care' may be inaccurate. For example, in the case of frailty, appropriate treatment setting would depend on social factors such as the family support that the patient has at home.

Data requirements for model 1
- List of clinically-proven treatments for every combination of condition and patient "type";
- Weightings for each element of treatment (which may or may not all be equal);
- Record of care that each patient has received;
- All diagnoses that each patient has received;
- Coverage of a sufficient timeframe which will be dependent on the condition.

The method could be adapted for use with aggregated data and consider population segments rather than individual patients.

Other considerations

- The threshold gap score would need to be defined according to the programme design and objectives. For example, it could be set to target a certain number or proportion of patients or it could be defined in absolute terms.
- Certain treatments may be considered as a higher priority than others and consequently, different weightings may be assigned to different treatments to calculate a weighted gap score rather than assigning equal weighting to each recommended treatment. Assigning weights to different treatments would require judgement and introduce an opportunity for bias.
- Investigation would be required to determine if those with the highest capacity to benefit are actually impactable.

Figure 6 below is an example of the NICE recommendations for diabetes (NICE, n.d.-b), which shows how the user is able to 'click through' the relevant parts of the flow chart to access the guidelines that are relevant for particular types of patients or circumstances.
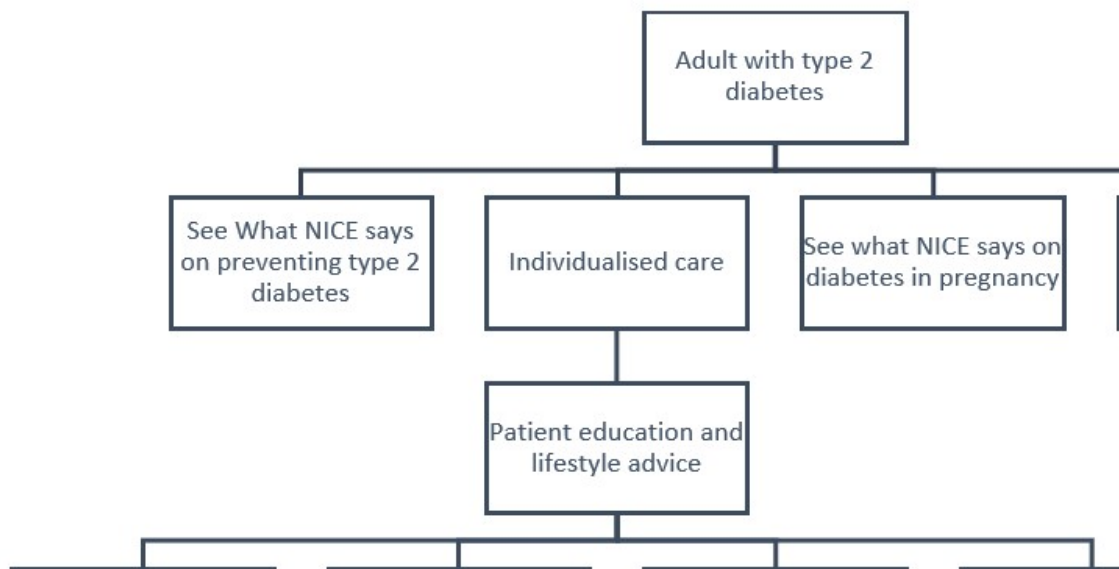

Figure 6: An overview of Type 2 diabetes in adults. Reproduced based on Nice Pathways (NICE, n.d.-b)

5.2.2 Impactability model 2: Condition severity benchmarking

Individuals or specific population segments can be benchmarked for a range of metrics covering risk characteristics of the individuals/segments, service utilisation and condition severity.

Method 2(a) involves identifying individuals/segments whose condition is more severe than would be expected for their given risk characteristics.

Method 2(b) involves identifying individuals/segments whose service utilisation is higher than would be expected for their given condition and/or condition severity.

Risk is not the same as condition severity. Risk is also not considered to be the same as outcomes. There may be circumstances where these items are defined to be the same thing e.g. The Electronic Frailty Index ("eFI") score can act as a measure of risk and as a measure of condition severity. However, this will not always be the case.

"Risk characteristics" must be clearly defined and will need to include the following elements:

- Derived from characteristics of an individual person or segment;
- that can be measured
    - within a timeframe that allows preventative interventions to be made
- which are not in themselves an indicator of current poor health;
- but which have a strong correlation with future adverse health outcomes;
    - "adverse health outcomes" needs to be defined. For example, unplanned hospital admission.

Statistical methods which can be applied within this impactability model are detailed further in Appendix 3.

Advantages and disadvantages for model 2
Advantages
- The impactability model could be built with aggregated population data rather than person-level data and so it could be tested in situations with limited data.
- Logical / rational (for analysts, but maybe less so for health professionals)
- Uses data that is readily available (EMRs, hospital records, community and social care provider records)
- If the aim is to reduce service utilisation or reduce risk or reduce condition severity then this method creates the metrics for measuring the success of itself.
- The method lends itself to statistical methods.

Disadvantages
- Many possibilities for defining the risk score, condition severity and service utilisation and there is not a single clear preferred choice for these metrics.
- Need clinical expertise to set the appropriate interventions for the range of values of each metric.

Data requirements for model 2
- Independent variables for calculating:
    - Risk score;
    - Condition severity; and
    - Service utilisation.
- All diagnoses that each patient has received;
- All health services that each patient has received;
- All factors that support the risk score;
- Over a sufficient timeframe

The method could be adapted for use with aggregated data only and consider population segments rather than individual patients.

5.2.3 Impactability model 3: Propensity to benefit scores

The propensity to benefit score is a method of prioritising high-risk individuals for treatment plans according to a score calculated for each individual. The scores are calculated using a predictive model that has been calibrated to health data using regression analysis. In one particular study, the approach was applied to insured members of a health plan who qualified for a high-risk case management programme (HRCM) (Hawkins et al., 2015). For any health condition with a specific preventive programme, a preventive programme may either benefit or be detrimental to the health of an individual. The benefit is a condition-specific measure and can be defined as a percentage. For example, for a patient with high cholesterol, the benefit of the preventive programme could be defined as the percentage improvement in their cholesterol level. For a patient with diabetes, the measured benefit of the preventive programme could be defined as the percentage improvement in their blood glucose level.

GLM (Generalised Linear Model) techniques can be applied to predict the benefit of a preventative programme for an individual before they join the programme. The main output of the predictive model is the **"propensity to benefit score"**, or **"propensity-to-succeed (PTS) score"** (as developed by Hawkins et al. (2015)). This method involved fitting logistic regression models to identify three groups of individuals: those most likely to engage in the preventative programme; those receiving high quality of care once engaged; and those incurring enough monetary savings related to programme participation to more than offset programme costs" (Hawkins et al., 2015). The independent variables in the regression analysis were membership details alongside medical and pharmacy claims data.

Advantages and disadvantages for model 3
Advantages
Preventive programmes can "contribute to a better quality of life by reducing the needs for radical treatments, such as surgery or chemotherapy" (Luijten, 2010). Preventive programmes such as "healthy nutrition & hydration, exercise/outdoor activity, social interaction, opportunities to engage in new activities to stimulate cognitive development" (Duffy, 2018) are generally beneficial for the health status of patients. This will help patients to move from one risk group to another risk group, which can directly inform investigation of the impactability of a preventive programme. A key advantage of the propensity to succeed method is simplicity and re-usability. Standard software can be applied for fitting the logistic regression models.

Disadvantages
The preventive programme may result in false positive results, which is a problem for impactability modelling. The false positive results might lead to wrong predictions of probabilities for propensity to benefit from a given preventive programme. According to Hawkin et al. (2015), the limitation of the propensity to succeed method is the quality of care dimension and quality of life measurement. An important issue to consider is the point at which the advantages of a preventive programme can overcome the disadvantages.

Data requirements for model 3
- All diagnoses that each patient has received
- All health services that each patient has received
- All factors that support the propensity-to-benefit score
- A measure of what "benefit" means for each condition and treatment
- Over a sufficient timeframe
- The method could be adapted for use with aggregated data only and consider population segments rather than individual patients.
- Specific preventive programme data

5.2.4 Impactability model 4: Patient activation measures

Patient activation describes the knowledge, skills and confidence a person has in managing their own health and care (NHS England, n.d.-a). This method for measuring impactability implicitly assumes a direct link between patient activation and the impact that an intervention will have on a given person. It is evident that the validity of this assumption will depend on the nature of the intervention - for example, the extent to which the intervention relies upon self-management by the patient.

The application of this model of impactability therefore depends on defining quantitative measures for:

 i.   Patient activation; and
 ii. The importance of self-management to the modelled intervention

There is a standard method for measuring patient activation in the form of the Patient Activation Measure (PAM), which the NHS uses as a commercially licensed tool from Insignia Health LLC. This involves assigning a PAM score between 0 and 100 to an individual at a given point in time, based on their answers to a survey consisting of 13 questions. The PAM score is then converted into one of four defined measures of activation, ranging from people who tend to be passive about their care and feel overwhelmed by managing their own health, to those who adopt many good behaviours to support their health (but may still struggle to maintain them when faced with life stressors) (NHS England, n.d.-a).

The existence of a score for patient activation enables interventions to be developed to meet two types of objective:

- Tailor interventions so that they allow for each individual's propensity to manage their own ongoing care and thereby optimise impactability for all patients. In this case, the PAM score (measure (i) above) is considered fixed, and the intervention itself is varied to optimise impactability fairly within a given funding envelope across the cohort of patients; this would be achieved by varying the importance of self-management to the intervention (ii). This approach would be expected to lead to directing more ongoing care resource towards those with lower PAM scores (reducing the need for self-management), while for high-PAM patients there would likely be more of an emphasis on providing guidance for self-management.

- Provide interventions with the express purpose of increasing a patient's PAM score, so that they become more empowered to manage their own ongoing care. These could involve education or peer support networks, for example. For such interventions, measuring impactability would involve identifying the patients most likely to experience an increase in their PAM score. A study by the King's Fund has suggested that patients who start with a low activation score are likely to experience the greatest increases to their scores from appropriately targeted interventions (Hibbard and Gilburt, 2014).

Insignia Health manages the process of converting survey responses to PAM scores and levels. This commercial inaccessibility of the full model means that PAM is not a suitable model for this Working Party to investigate in detail.

However, the concept of patient behaviour measures is a useful comparative reference to ensure a broad evaluation of the methods that are taken forward for more detailed analysis.

Advantages and disadvantages for model 4
Advantages:
- Tends to direct more resources towards cohorts who are less ready to manage their own ongoing care, thereby helping to address inequality.
- It could help to reduce avoidable use of healthcare services, for example by providing appropriate self-management guidance to those with higher activation scores, rather than unnecessary additional care.
- Encourages the development of interventions that are more personalised to individual patients' level of self-management readiness, including providing patients with more choice and control over the way their care is planned and delivered to them.
- PAM is an existing, recognised scoring mechanism for patient activation.

Disadvantages:
- Not suitable for detailed quantitative analysis by this Working Party, because the scoring mechanism is commercially owned.
- Reliant on collecting survey responses, so in practice could only be applied to targeted cohorts, pathways or interventions.
- Surveys can also be vulnerable to natural fluctuations in the subjective interpretation of the questions between different patients.

Data requirements for model 4
- Requires patient survey data, which would likely be more time consuming and expensive to collect (compared to patient data sitting in Electronic Medical Records (EMRs) and hospital systems, for example) and would be available for only a fraction of the population.
- Weightings for each element of activation/engagement (which may or may not all be equal)
- To use patient activation as a measure of impactability, there would be a need to define and quantify how important self-management is to the success of a given intervention

5.3 Summary

From an initial list of impactability models, four were identified for a more thorough review. A summary of the pros and cons of these four approaches, as well as the key data requirements, is shown below.

	Pros	Cons	Data Required
Gaps-In-Care Scores	Based on proven clinical theory Encourages consistency in prescribed treatments Ability to measure the effectiveness of implementation Likely to reduce health inequalities	Requires agreed list of recommended care/treatment/intervention guidelines for each condition and patient group. Challenge to set appropriate weights for each element of care.	List of clinically-proven treatments for every combination of condition and patient "type" Record of care that each patient has received All diagnoses that each patient has received
Condition Severity Benchmarking	Uses data that is readily available (EMRs, hospital records, community and social care provider records) If the aim is to reduce service utilisation or reduce risk or reduce condition severity, then this method creates the metrics for measuring the success of itself.	Many possibilities for defining the risk score, condition severity and service utilisation and there is not a single clear preferred choice for these metrics. Need clinical expertise to set the appropriate interventions for the range of values of each metric.	All diagnoses and health services that each patient has received All factors that support the risk score Model could be built with aggregated data only rather than individual patients.
Propensity-to-Benefit Scores	The method lends itself to statistical methods, which actuaries can deploy. Standard software can be applied for fitting the logistic regression models.	Many possibilities for defining the propensity-to-benefit score and there is not a single clear preferred choice for this metric. Questionable accuracy of results, such as false positives.	All diagnoses and health services that each patient has received All factors that support the propensity-to-benefit score Specific preventive programme data A measure of what "benefit" means for each condition and treatment
Patient-Activation Scores	Helps to reduce avoidable use of healthcare services Encourages more personalised interventions Likely to reduce health inequalities	Data not readily available across whole populations Can only be practically deployed for small segments of the population, which must be first identified by some other method. Data vulnerable to subjective responses of patients. Existing versions of scores are commercially-owned so not accessible for study by this Working Party.	Patient survey data Weightings for each element of activation/engagement need to define and quantify how important self-management is to the success of a given intervention.

Table 1: The pros, cons and data requirements of selected impactability models

Section 6: Ethics, Patients and the Public

Lead authors of Section 6: S. Culkin, M. Chan, S. McKeigue, J. Robertson, J. Umpleby.

This section aims to support consideration of ethical issues, as well as considering patient and public interests. This is a wide-reaching area for consideration including privacy in terms of data governance, patient involvement and health inequalities, to name a few. This section provides an overview of the work undertaken by the working party so far in this area and should therefore, not be read as exhaustive guidance but a starting position for required discussions.

It is important to consider the ethical implications of proposed changes that concern the health of the population. Ethics is the setting of ideals to help guide decisions and actions and overcome biases, to make decisions more fairly, objectively, or scientifically (Spike, 2018). This could be considered as formally considering what we may instinctively feel is right or wrong. In the context of population health management, it involves a process of deliberation; individuals and groups voice differing views on relevant values and the respective weighting of those values. This is done with the aim of working toward a consensus but importantly acknowledging where there may be continued disagreement.

The objective of this section is to set out a deliberative approach to ethics, i.e. create a structure for how ethical issues could be considered for impactability modelling, thereby supporting this new area of health research. When modelling and making decisions, assumptions must be made about the real-world. This first step in the process could be to consider the ethical implications of the data that is used in making those assumptions, used in the models, and so informs decision-making. In future ethical deliberations, those using impactability modelling should be cautious to avoid using ethical discussions where there is actually a need for the use or creation of regulations (this is known as ethics washing) (Wagner, 2018).

In this section, ethics principles, ethical frameworks and citizen involvement are explored. We have split the ethical considerations into two sections. The first looks at the ethical considerations of impactability model inputs i.e. the ethical use of health data (section 6.1). The second section (section 6.2) looks at the ethical considerations of impactability model outputs which aligns well with existing public or population health ethical considerations. In this section, we also consider the potential pitfalls of bias or unintentional deepening of health inequalities. To help practitioners surface and record underlying beliefs and values we introduce a method to become more aware and so potentially address these pitfalls. Finally, in section 6.3, we look at the importance of involving patients and the public in impactability modelling development and use.

Those working in this area may need to seek advice regarding compliance and appropriateness in accordance with laws and regulations that are specific to the context they operate within.

6.1 Ethical considerations for model inputs e.g. data use

A key component for the ethical consideration of impactability modelling is the data used in that modelling. Today there is more data and more types of data generated than ever before. In addition, there is a wider range of analytic and data science techniques possible that can turn data into actionable insights. This actionable insight can allow for improved strategic planning and assist clinical decision-making. However, the use of this data, in general and for impactability modelling, can potentially raise a number of ethical issues and as such these need to be considered.

There are existing data ethics principles that can be leveraged. For the purpose of impactability modelling we have summarised these in to 3 areas: what is the need, what data and how the data is used.

The Department of Health and Social Care and NHS-England has developed a code of conduct to enable the development and adoption of safe, ethical and effective data-driven health and care technologies (United Kingdom Government, 2018a) [1].

The Department for Culture, Media and Sport have also produced a data ethics framework to set out clear principles for how data should be used in the public sector (United Kingdom Government, 2018b) [2].

Together they provide a set of principles for the use of data in impactability modelling.[4]

Based on the two published Government data ethics frameworks [1, 2] we propose a list of the key principles that should be followed when carrying out impactability modelling, grouped in 3 sections *what is the need, what data, and how the data is used.*

What is the need?
Be clear, from the start, on what the expected benefit of the modelling is for patients or your population.

Before starting to work with data you must consider the user need and expected benefit to patients or your population. Having a clear sense of user need from the beginning will help you clarify the problem you are trying to solve, even if you don't yet know what the solution is, or even the path to the solution.

Who is the data subject?
Be clear on whose data will be used as a model input and why.

In some cases, the user, data subject and beneficiary of the modelling may be the same person. If they're not, make sure you understand the different interests and expectations of each. In particular, you need to be careful you don't use some people merely as a means to benefit others. Understanding this is fundamental to making sure you are treating people properly.

See Data Ethics Framework Principle 1 and Code of Conduct Principle 1 for further details.

What data?
Use data that is proportionate to the need and is in line with any relevant guidelines on its purpose.

Data collection, storage or analysis must be proportionate. Proportionality will depend on the user need and benefit, whether the data is personal (and, if so, whether it can be de-identified), and the source of the data and how it was collected. If the proposed data use isn't proportionate to the need then an alteration to the sources, collection methods or analysis should be considered - for example by considering if there is some other way to meet the need. A good rule of thumb here is to consider if you could justify why the data was needed, how it is meeting the need and so why it was sourced, to a member of the public. If this is harder to justify, it may not be an ethical use of the data. Taking time to articulate the relationship between the data used as an input to impactability models and the conclusion that can be drawn from the results of that model, should also help you defend your use of the data in that model at a future date to stakeholders including management and members of the public.

---

[4] The Institute and Faculty of Actuaries and Royal Statistical Society also published a guide for ethical data science in 2019,
Institute and Faculty of Actuaries; Royal Statistical Society, 2019. *A guide for ethical data science: a collaboration between the Institute and Faculty of Actuaries (IFoA) and the Royal Statistical Society (RSS)*. https://www.actuaries.org.uk/documents/guide-ethical-data-science

The following questions will help you determine whether your data is proportionate to the modelling purpose:

- Is the measure suitable to achieve the aim of the modelling?
- Is the measure necessary to achieve the aim of the modelling?
- Would the proposed use of the data be deemed inappropriate by those who provided the data or other stakeholders?
- Would any proposed secondary use of the data make it less likely that stakeholders would want to give access to that data for the primary purpose again? i.e. would there be any regrets?

See Data Ethics Framework Principle 3 and Code of Practice Principle 3 for further details.

Be fair, transparent and accountable about what data is being used.

When developing your modelling, you need to make sure you are compliant with all relevant laws and codes of practice. This includes, but it is not limited to, transparency of use on profiling and automated decision-making that is intended. The working party recommend that appropriate legal advice is sought prior to data collection or processing to ensure compliance and appropriateness of action throughout. This recommendation applies within a UK context or elsewhere.

For a UK context, some important pieces of legislation that typically apply are:

EU General Data Protection Regulation (GDPR)

Since May 2018 the national data opt-out allows people to opt out of their confidential patient information being used for purposes beyond their individual care and treatment. By 2020 any health and care organisation that processes and/or disseminates data that originates with the health and adult social care system in England is required to be in compliance with the national data opt-out policy. Anonymised data in line with the ICO's code of practice is exempt from this.

Data Protection Act 2018

The Data Protection Act controls how personal information is used by organisations or the Government. Everyone responsible for using personal data has to follow strict data protection principles. There is stronger legal protection for more sensitive information and individuals have the right to find out what information organisations store about them.

These requirements are particularly strict when processing health data, which is considered a special category data under the Act. The Act introduced new obligations that require integration of data protection concerns into every aspect of processing activities. This approach is known as 'data protection by design and by default'. From a practical perspective, the important documents underpinning this are data flow maps, data protection impact assessments and privacy notices.

For example, under the Data Protection Act 2018, individuals have the right to be informed about the collection and use of their personal data. A privacy notice should identify who the data controller is, with contact details for its data protection officer. It should also explain the purposes for which personal data is collected and used, how the data is used and disclosed, how long it is kept and the controller's legal basis for processing.

In addition, the NHS has a number of safeguards in place to assure patients that their data is managed safely and securely, and their rights to privacy and confidentiality are upheld. These include the NHS Constitution and the Data Sharing and Privacy Toolkit.

Data protection-by-design principles can help you abide by these protocols, including the development of data-sharing agreements, data flow maps and data protection impact assessments.

See Data Ethics Framework Principle 2 and Code of Practice Principle 4 for further details.

How the data is used?
Understand and be transparent about limitations

Even when legal and proportionate, limitations to the data can make a proposed approach inappropriate, unreliable or misleading.

First, data should be reviewed for *accuracy* and *completeness*; where accuracy is the closeness of agreement between a data value and its true value, and completeness is the presence of the necessary data.

Second, consideration should be given to the representativeness of the data used for developing the model and how well it reflects its target population. If the data provided for development is limited to certain demographic categories or disease areas, this could potentially limit the applicability of the model in practice as its ability to accurately predict could be different in other groups.

Third, any biases in existing datasets should be identified and controlled for otherwise there is a risk of these being embedded and obscured from accountability. Bias in datasets can be introduced in various ways, including collection techniques, limited representativeness of a cohort and social bias from historical decision making. Accounting for bias is vital in ensuring that any gathered evidence does not inadvertently produce discriminatory decisions and that there is not the possibility of a single data predictor having an inflexible impact on the forecasted outcome.

The principle to 'understand and be transparent about limitations' discusses the presence of bias. One potential way of how biases may inadvertently be included in impactability modelling is due to the large difference in the quantity of readily accessible care data in different setting i.e. hospital activity data may be comprehensive, but community delivered health or social care data may not be. This is also an issue when we consider the availability of information on health outcomes. For example, through hospital data there is information on emergency admissions that could potentially be avoided, but less routine data on health and wellbeing outcomes. The consequence of this could be unconsciously to focus modelling efforts on areas where data is most abundant or easy to access instead of the other variables that may be of interest, e.g. on quantity or cost of hospital care rather than health and wellbeing outcomes. Proponents of impactability modelling should be aware of this particular scope for bias.

Fourth, investigation should be made concerning the sensitivity of the results to the method of handling missing values. Missing data poses a threat to the quality of the model and handling methods such as imputation techniques which attempt to compensate for missing values can be prone to bias. Therefore, the implications of working with the data in this way should also be considered fully.

Finally, assessment of data quality should not be a one-off check but continuous. NHS England and the UK Statistics Authority have produced guidance on data quality, which should be referred to. impactability modelling should adapt to the changing context which implies a need for continuous consideration and scrutiny of data use.

See Data Ethics Framework Principle 4 and Code of Practice Principle 6 for further details.

Make use of open standards

It is considered worthwhile to utilise and build into the model the current data and interoperability standards available to ensure it can communicate easily with the existing national systems. Programmatically building data quality evaluation into impactability modelling development helps to ensure that harm does not occur due to poor data quality creeping into the process at a later stage.

Within the health and social care system, information standards cover the specifications used to collect and extract data from information technology systems. NHS Digital currently hosts a range of data, clinical and interoperability standards for the health and social care network which are believed to be fruitful in this area.

See Code of Practice Principle 5 for further details.

Make security integral to the design

By safeguarding data and integrating appropriate levels of security into the design of devices, applications and systems, and keeping in mind relevant standards and guidance it is possible to keep data safe.

A core element of adoption is to ensure that security and data protection methodology have been incorporated. NHS Digital has launched a new Data Security and Protection Toolkit to ensure that patient information is kept safe. All organisations that have access to NHS patient data and systems must complete the toolkit to provide assurance that they are practising good data security and that personal information is handled appropriately.

See Code of Practice Principle 9 for further details.

Summary
What needs?
Are you clear about the aim of the impactability modelling and the expected benefit to your patients or population at the start?

What data?
Are you collecting, storing and using only data that is suitable and necessary to achieve your modelling aim?

Are you using data for a purpose for which it was collected?

Have you completed the relevant data sharing agreements and privacy impact assessments?

How is it used?
Do you understand and are transparent about the limitations and biases in the data?

Have you made full use of open standards?

Have you made security of the system integral to the design?

Those working in this area may need to seek advice regarding compliance and appropriateness in accordance with laws and regulations that are specific to the context they operate within.

6.2 Ethical considerations for model outputs e.g. population health ethics

Ethical Framework

From reviewing various relevant ethics frameworks, we have provided a proposed ethical framework for impactability modelling. This is an area that can continue to be developed and improved as it is put in to practice.

This framework should not be seen solely as a box ticking exercise and should be used proportionately for the specific modelling exercise. It is best used at the point where information on the model framework and planned intervention is available, but before model and intervention design is confirmed, to allow for any modifications. The possibility of increasing ethics support available to users should be explored.

Population health ethics stems from traditional medical ethics, which places the principles of beneficence ('do good') and non-maleficence ('do no harm'), at the forefront, and also brings in the principles of autonomy and fairness, without order or rank (Spike, 2018). A scoping review of proposed population health ethics frameworks showed that these four principles formed the core of all frameworks. Further information on these four core principles can be found in Appendix 4.

Proposed ethical framework for impactability modelling

The scoping review identified four studies which recommended population health ethics or health modelling frameworks (Boden and McKendrick, 2017, Fourie et al., 2014, Coggon and Viens, 2017, Tannahill, 2008). We have produced an ethics framework for impactability modelling, based on these studies, see Table 6. This framework is given as an example of one that could be used to inform how the model and its outputs are used.

Principle	Impactability modelling-specific factors	Evaluation of ethical risk
1. Utility	Cost-benefit and effectiveness, practicality	1. How is utility assessed in the model and in decision making processes? 2. Are alternative assessments feasible, useful, and likely to give similar results? 3. How likely are interventions based on model predictions to be implemented, given the target population and context, model assumptions and uncertainty, and political will?
2. Producing benefits	Benefit vs harm	1. If a policy decision is based on model evidence, is it more likely to be effective, beneficial and safer than a decision made in the absence of the model? 2. Has the model been verified, and validated in the target population and context? 3. Have potential harms in each subpopulation been identified and modelled/factored into the decision-making process? 4. What are the sources and magnitude of uncertainty—are these associated with parameter uncertainty, model assumptions and/or model selection?
3. Fairness	Modelling procedure and distribution of resources	1. Is any lack of knowledge about important parameters attributable to uncertainty, variability or lack of data? 2. Is lack of data or parameter uncertainty related to hard-to-reach subpopulations or choice of outcomes? 3. Is model variability attributed to known factors, to create more ethical outcomes? 4. If interventions based on model predictions are implemented, can the predicted benefits and harms to different individuals and subpopulations be quantified?
4. Procedural justice	All modelling and deliberative procedures are rigorous and adhered to	1. Are all the relevant procedures set up by each stakeholder followed and documented? 2. Are procedures rigorous, leading to a high quality modelling and decision making process? 3. Are procedures relevant to this context and set up prior to model development and decision making? 4. Have relevant stakeholders been actively consulted and feedback considered and acted on?
5. Transparency	Transparency of procedures and assumptions	1. Are model development process and assumptions well documented and available for scrutiny? If permitted, are they made public? 2. Is there information on potential conflicts of interest, constraints, or biases affecting data collection and analyses? 3. Are model documentation and results understood by stakeholders?
6. Independence	Scientific independence in modelling, patient autonomy	1. Is the model provenance known and well documented (e.g. funding sources, conceptual design, coding, verification, review processes and publication, as well as the modellers involved)? 2. Has the model been validated using independent data sources not used in its parameterisation? 3. Do interventions based on model predictions restrict patient autonomy or patient choice? Are individual circumstances or preferences considered?

Table 6: Proposed ethical framework for impactability modelling, synthesised from existing frameworks
(Boden and McKendrick, 2017, Fourie et al., 2014, Coggon and Viens, 2017, Tannahill, 2008)

Summary - What to do next:

- Increase awareness of ethical considerations related to impactability modelling:
  - Ethics as a process of deliberation that guides decision making to be more fair, objective, or scientific
  - The 4 main population health ethics principles: beneficence (do good), non-maleficence (do no harm), fairness and autonomy
  - Additional principles such as utility, transparency and procedural justice
- Factor in other real-world considerations for model output evaluation and implementation
- Actively assess fairness in selection of interventions, and how these processes may affect health inequalities
- Incorporate the proposed ethical framework for impactability modelling in Table 6 into the modelling process
- Consider refinements to the modelling process and the ethical framework

Box 4: Case Study: Lessons from Policing

Similarly to health care, in recent years there has been an increase in to use of algorithmic tools in policing with the promise to improve the police's decision making and prediction abilities. One of the first tools to be used by a UK police force was the Durham Constabulary's Harm Assessment Risk Tool (HART). The tool was developed to aid decision making when assessing the risk of future offending to enable greater and more confidence in the use of 'out of court disposals'. Using this tool as a case study the concept of 'experimental' proportionality, and the ALGO-CARE guidance framework have been proposed (Oswald et al., 2018).
ALGO-CARE is proposed as a guidance framework covering the key legal and practical concerns that should be considered when using algorithmic risk tools by the police (Table 7).

A	Advisory	In general the assessment from a tool should only be seen as advisory and the human officer should retain decision making discretion.
L	Lawful	Multi-faceted point: Is there a clear policing purpose to justify the use of an algorithm? Is the potential interference with the privacy of individuals necessary and proportionate? Can the tool be demonstrated to improve the current system? Is data lawfully obtained, processed and retained? Is the operation of the tool compliant with national guidance?
G	Granularity	Is the tool at the appropriate level of detail for the purpose? Are technological, data quality and data provenance issues understood and appraised?
O	Ownership	[Relevant if the tool is developed by external partners] Who owns the tool and source code and how will it be updated?
C	Challengeable	Individuals should be notified of the use of algorithmic tools. Regular validation and recalibration should be based on publicly available rules.
A	Accuracy	Can the predictive power be justified and the consequences of false/inaccurate predictions be explained
R	Responsible	Ethical considerations are factored into the use of tools and it is recommended that an ethical review committee is established.
E	Explainable	Can the algorithm be explained and justified?

Table 7: ALGO-CARE guidance framework

The elements in the framework are broadly consistent with the principles included in the code of conduct and data ethics framework. However, the framework does raise some (if not substantial) insights.

The ownership element raises the ethical consideration on when a tool is developed by an external contractor. This could be the case for impactability modelling and especially Population Health Analysis more broadly. There is the focus on who owns the algorithm and data analysed, and from an ethical perspective there are the points as to (i) whether the contractor is adhering to data ethics principles in their collecting, processing, handling of data, (ii) whether the ownership of the model allows adequate transparency on the details of the algorithm and how it is used, and whether the ownership allows updating of the model so using the latest data to maintain relevancy.

The accuracy element proposes the point that the consequences of inaccurate outputs/predictions should be able to be explained, this would be sensible to also apply to impactability modelling. The advisory element could also be directly translated to ethical principles for impactability modelling that (the algorithmic part) of impactability modelling should only be used in an advisory capacity and the health care professional retains decision making discretion.

Fairness and inequalities: A method for capturing values and beliefs - reflexivity exercise

As mentioned above, the concept of fairness (see, principle 3) (can be considered in terms of health inequalities) is particularly important ethical consideration when looking to model or predict the impact of different interventions, because it can influence whether the intervention is recommended or not.

When building, choosing, parameterising and using a model there are many choices faced by the user. The decision-making process is informed by prior experiences, values and beliefs. These elements can either be consciously or unconsciously embedded into a model, and could lead to unfairness and greater health inequalities.

Therefore, a 'reflexivity exercise' has been developed and piloted. It aims to bring awareness to our prior experiences, values and belief that frame our actions and decisions.

The consequences desired and undesired of impactability modelling, and of wider PHM, was considered by each member individually and openly discussed as a group. This created an opportunity for the working party to enhance its understanding and expand its ethical consideration. It was considered important to take time to reflect on the nature of our own views and how they may differ from others.

Reflexivity is "the conscious examination of past experiences, thoughts and ways of doing things" (University of Edinburgh, 2019). This was considered an important element of model development in a health context as it "challenges the status quo of practice, thoughts and assumptions" (University of Edinburgh, 2019). It "may therefore inform our decisions, actions, attitudes, beliefs and understanding about ourselves" (University of Edinburgh, 2019).

An initial attempt at a self- and group-reflection exercise was piloted by the working party and an example, along with practical guidance, is provided in Appendix 5 so the reader may also undertake the process.

This process seeks to:
- Increase/improve performance and skills;
- Increase awareness of ability and attribute and provide evidence for these;
- Evaluate the quality and success of action plans; and,
- Apply theoretical knowledge/frameworks to real experiences to expand understanding of underlying theory.

The following learnings resulted from the pilot:
- Provide sufficient time for reflection for both self (minimum of 1.5 hours) and group (minimum of 2.5 hours) aspects as allowing time for discussions to naturally unfold creates fruitful learning opportunities;
- Provide a format where all in the group feel comfortable contributing to discussions with consideration for group size and independent/experienced facilitators
- A wide range of valid views highlighted the complexity within health modelling and the need for clarity and transparency;
- A structured process was helpful in navigating reflection within a group environment.

It is recommended that others undertaking impactability modelling take part in this or a similar exercise to reflect on and capture beliefs and values that may influence model assumptions and parameters.

Summary - What to do next
- Create time early in the project management of model development to discuss reflection, the appetite to undertake work in this area and agree a framework to undertaken reflection
- Discuss how reflection may benefit the outcome of the project, create a plan and allocate time accordingly
- Pilot a self-reflection exercise ensuring that actions and improvement are the key outcome
- Pilot a group reflection exercise with a facilitator, record a summary of discussions and actions and improvements, create a plan of how actions and improvement will be further embedded into work

6.3 Patient and Public Engagement

Section 6.2 provided guidance on ethics for those involved in impactability modelling. This section moves on to another important area for consideration; how to engage patients and the public.

Why is public and patient involvement important?
Involvement is vital to check that your approach is:
- Grounded in the reality of the condition(s) being considered;
- Understood by patients and the public (e.g. not 'black box' or a computer generating an arbitrary result); and
- In line with reasonable expectations of citizens.

What kind of areas should you cover?
The following areas are a good starting point, not an exhaustive list. You may wish to explore them with patients and the public, to get their insight into:

The data being used:
a. In general, is the data being used in line with reasonable expectations of citizens?
b. If you are planning on using **wider determinant data and other data sources not directly from the health system** do patients and the public feel the data is being used appropriate? Would they reasonably expect the data to be used in this way?
c. Are you are considering using data *created by* patients and the public (for example completing **questionnaires** about their health and life,) rather than data *collected about* them? If so, consider gathering ideas about what questions to include in these questionnaires
d. Collect ideas and suggestions about other sources of data to consider

Model design:
a. Can you glean insight from patients about their condition and their life, to help with model design, for example through mapping out a typical patient pathway or drawing out a flow diagram to show lifestyle interactions impacting on a condition?
b. When you are designing your model, are different model types (see Section 5) better understood or accepted?

Model use:
a. Once your model is ready for use, test it out with patients and the public:
    i. Is the model answered a question which is important to the patient?
    ii. Do they understand how the model is generating results?

   iii. And how do the public and patients feel about how the results are used, for example are they more comfortable with outputs being used to provide decision support rather than fully automating decisions?

 b. Also consider involving patients and the public in sensitivity testing to see how outputs change and inputs are varied.

 c. Explainability and transparency are crucial - try out different ways and see which are better accepted

 d. And include the steps above as part of your ongoing evaluation of the impactability model

General considerations for involvement

There is a Consensus Statement on Public Involvement and Engagement with Data Intensive Health Research which contains a lot of useful information to guide involvement (Aitken et al., 2019). It is based around the following 8 key principles for ensuring inclusion of the public in data-intensive health research:

1. Have institutional buy-in
2. Have clarity of purpose
3. Be transparent
4. Involve two-way communication
5. Be inclusive and accessible to broad public
6. Be ongoing
7. Be designed to produce impact
8. Be evaluated

The NHS England Participation Hub contains a wide range of resources which may be useful in this area. Some of the key guides are listed below:

Planning for participation
https://www.england.nhs.uk/wp-content/uploads/2014/03/bs-guide-plann-part1.pdf
Budgeting for participation
https://www.england.nhs.uk/wp-content/uploads/2014/03/bs-guide-budget-part.pdf
Online participation
https://www.england.nhs.uk/wp-content/uploads/2016/07/bitesize-guide-online-participation.pdf
Focus groups for participation:
https://www.england.nhs.uk/wp-content/uploads/2016/07/bitesize-guide-focus-groups.pdf

Further practical considerations for involvement:

There are a range of options for patient and public engagement, with different resource implications, as shown in the figure below:
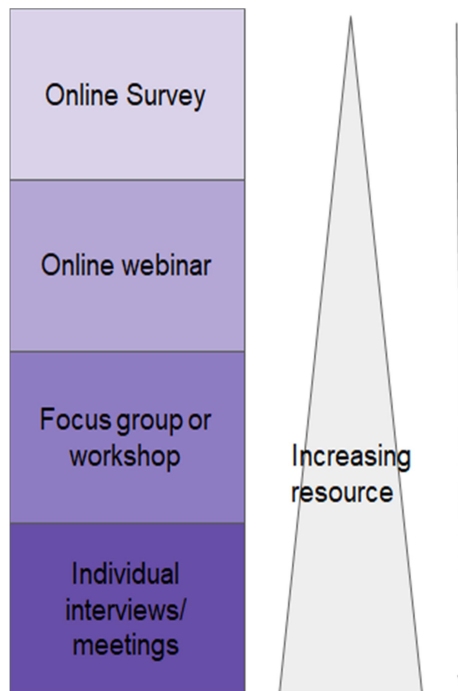

Figure 7: Diagram showing spectrum of participation methods

There are advantages and disadvantages to the different methods - the most appropriate approach(es) should be selected based on what you are trying to achieve and the budget and time you have. For example, an online survey with the option to join a follow up webinar can work well (but be aware of data protection considerations if any personal information such as email address is collected).

There are a range of useful online tools that are often free.

If you would like to meet face to face with participants, some practical things to consider are:

- How do you recruit to a role? - does it need to be advertised with a role description and an estimate of commitment, travel requirements etc.
- How will you pay for any expenses incurred by participants?
- Requirements for safeguarding and accessibility

Other considerations include:

- How can you ensure a representative sample? or that different types of patient or member of the public are involved, e.g. newly diagnosed vs. a long term expert
- How will you use and save any personal stories or other information they may share with consideration for relevant data governance?

Summary - What to do next:
- Engage patients and the public on:
  - The data being used
  - How the model is designed and developed
  - How the model outputs are used
- When engaging, consider the points made in the Consensus Statement on Public Involvement and Engagement with Data Intensive Health Research (Aitken et al., 2019)

References

AITKEN, M., TULLY, M. P., PORTEOUS, C., DENEGRI, S., CUNNINGHAM-BURLEY, S., BANNER, N., BLACK, C., BURGESS, M., CROSS, L. & VAN DELDEN, J. 2019. Consensus statement on public involvement and engagement with data-intensive health research. *International Journal of Population Data Science,* 4. https://ijpds.org/article/view/586

BARTLEY, M. 2016. *Health inequality: an introduction to concepts, theories and methods*, John Wiley & Sons.

BODEN, L. A. & MCKENDRICK, I. J. 2017. Model-based policymaking: A framework to promote ethical "Good Practice" in mathematical modeling for public health policymaking. *Frontiers in Public Health,* 5**,** 68. https://doi.org/10.3389/fpubh.2017.00068

BRAITHWAITE, R. S., STEVENS, E. R. & CAPLAN, A. 2016. Is risk stratification ever the same as 'profiling'? *Journal of Medical Ethics,* 42**,** 325-329. https://doi.org/10.1136/medethics-2015-103047

BUJA, A., RIVERA, M., SOATTIN, M., CORTI, M. C., AVOSSA, F., SCHIEVANO, E., RIGON, S., BALDO, V., BOCCUZZO, G. & DAMIANI, G. 2019. Impactibility model for population health management in high-cost elderly heart failure patients: A capture method using the ACG system. *Population Health Management*. https://doi.org/10.1089/pop.2018.0190

CLEGG, A., BATES, C., YOUNG, J., RYAN, R., NICHOLS, L., TEALE, E., MOHAMMED, M., PARRY, J. & MARSHALL, T. 2017. 129Development, Internal Validation And Independent External Validation Of An Electronic Frailty Index Using Routine Primary Care Electronic Health Record Data. Age and Ageing, 46, i35-i38. https://doi.org/10.1093/ageing/afx068.129

COGGON, J. & VIENS, A. M. 2017. *Public health ethics in practice: A background paper on public health ethics for the UK PHSKF* [Online]. Public Health England. Available: https://www.gov.uk/government/publications/public-health-ethics-in-practice [Accessed 1 June 2019].

COURT, S. D. 1981. Inequalities in health: Report of a research working group. *Archives of Disease in Childhood,* 56**,** 161-162. https://doi.org/10.1136/adc.56.3.161

DALAL, A., MORGAN, L., ROUX, C. & SHARMA, S. 2018. Make change happen: Getting insurance providers ready to better serve low-income households. *International Labour Oranization,* 2018**,** 37. http://www.impactinsurance.org/sites/default/files/IIWP53_0.pdf

DUBARD, C. A. & JACKSON, C. T. 2018. Active redesign of a medicaid care management strategy for greater return on investment: Predicting impactability. *Population Health Management,* 21**,** 102-109. https://doi.org/10.1089/pop.2017.0122

DUFFY, J. 2018. *Frailty: the new epidemic and how to manage it* [Online]. Royal College of Physicians. Available: https://www.rcplondon.ac.uk/file/9743 [Accessed 1 June 2018].

DUNCAN, I. 2004. Dictionary of disease management terminology. *Washington, DC: Disease Management Association of America*.

DUNN, M. & HOPE, T. 2018. *Medical ethics: a very short introduction*, Oxford University Press.

ELWELL-SUTTON, T., MARSHALL, L., BIBBY, J. & VOLMERT, A. 2019. Briefing: Reframing the conversation on the social determinants of health. The Health Foundation. https://www.health.org.uk/sites/default/files/upload/publications/2019/Reframing%20the%20conversation%20on%20social%20determinants.pdf

FADEN, R. R., KASS, N. E., GOODMAN, S. N., PRONOVOST, P., TUNIS, S. & BEAUCHAMP, T. L. 2013. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. *Hastings Center Report,* 43**,** S16-S27. https://doi.org/10.1002/hast.134

FOURIE, C., BILLER-ANDORNO, N. & WILD, V. 2014. Systematically evaluating the impact of diagnosis-related groups (DRGs) on health care delivery: a matrix of ethical implications. *Health Policy,* 115**,** 157-164. https://doi.org/10.1016/j.healthpol.2013.11.014

FREUND, T., WENSING, M., GEISSLER, S., PETERS-KLIMM, F., MAHLER, C., BOYD, C. M. & SZECSENYI, J. 2012. Primary care physicians' experiences with case finding for practice-based care management. *The American Journal of Managed Care,* 18**,** e155. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3422074/

GALTON, F. 1886. Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland,* 15**,** 246-263. https://doi.org/10.2307/2841583

HAWKINS, K., OZMINKOWSKI, R. J., MUJAHID, A., WELLS, T. S., BHATTARAI, G. R., WANG, S., HOMMER, C. E., HUANG, J., MIGLIORI, R. J. & YEH, C. S. 2015. Propensity to succeed: Prioritizing individuals most likely to benefit from care coordination. *Population Health Management,* 18**,** 402-411. https://doi.org/10.1089/pop.2014.0121

HIBBARD, J. & GILBURT, H. 2014. Supporting people to manage their health:. *An introduction to patient activation. London: The King's Fund*. https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/supporting-people-manage-health-patient-activation-may14.pdf

HIPPISLEY-COX, J. & COUPLAND, C. 2013. Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. *British Medical Journal open,* 3**,** e003482. https://doi.org/10.1136/bmjopen-2013-003482

JOHN HOPKINS ACG 2018. White Paper: The John Hopkins ACG® System. Johns Hopkins Center for Population Health Information Technology.

KINGS' FUND. 2018. *A vision for population health* [Online]. London: The Kings' Fund. Available: https://www.kingsfund.org.uk/publications/vision-population-health [Accessed 1 May 2019].

KNEALE, D., ROJAS-GARCÍA, A. & THOMAS, J. 2018. Exploring the importance of evidence in local health and wellbeing strategies. *Journal of Public Health,* 40**,** i13-i23. https://doi.org/10.1093/pubmed/fdx152

LEWIS, G. 2015. *Next steps for risk stratification in the NHS* [Online]. NHS England. Available: https://www.england.nhs.uk/wp-content/uploads/2015/01/nxt-steps-risk-strat-glewis.pdf [Accessed 1 June 2019].

LEWIS, G. H. 2010. "Impactibility Models": Identifying the subgroup of high☐risk patients most amenable to hospital☐avoidance programs. *The Milbank Quarterly,* 88**,** 240-255. https://doi.org/10.1111/j.1468-0009.2010.00597.x

LOVE-KOH, J., GRIFFIN, S., KATAIKA, E., REVILL, P., SIBANDZE, S. & WALKER, S. M. 2019. Incorporating concerns for equity into health resource allocation: A guide for practitioners. *Working Papers, Centre for Health Economics, University of York*. http://eprints.whiterose.ac.uk/147414/

LUIJTEN, N. 2010. *When do the advantages of preventative healthcare overcome the disadvantages?* Business Studies, Universiteit van Tilburg. http://arno.uvt.nl

MARMOT, M., ALLEN, J., GOLDBLATT, P., BOYCE, T., MCNEISH, D., GRADY, M. & GEDDES, I. 2010. The Marmot review: Fair society, healthy lives:. *The Strategic Review of Health Inequalities in England Post-2010*. http://cfoaservices.co.uk/pdf/ProfessorSirMichaelMarmot.pdf

MAY, T., BYONANEBYE, J. & MEURER, J. 2017. The ethics of population health management: Collapsing the traditional boundary between patient care and public health. *Population Health Management,* 20**,** 167-169. https://doi.org/10.1089/pop.2016.0058

NHS ENGLAND. 2019. *The NHS Long Term Plan* [Online]. NHS England. Available: https://www.longtermplan.nhs.uk/ [Accessed 1 October 2019].

NHS ENGLAND. n.d.-a. *Patient activation and PAM FAQs* [Online]. NHS England. Available: https://www.england.nhs.uk/personalisedcare/supported-self-management/patient-activation/pa-faqs/ [Accessed 1 June 2019].

NHS ENGLAND. n.d.-b. *Reducing health inequalities resources* [Online]. NHS England. Available: https://www.england.nhs.uk/about/equality/equality-hub/resources/ [Accessed 30 July 2019].

NHS ENGLAND, NHS DIGITAL, NHS IMPROVEMENT & PUBLIC HEALTH ENGLAND 2019. Population Health Management. The Association of Directors of Public Health (ADPH). https://www.adph.org.uk/wp-content/uploads/2019/02/NHS-England-Presentation.pdf

NHS ENGLAND, NHS DIGITAL & PUBLIC HEALTH ENGLAND. 2018. *PHM Flatpack. A guide to starting PHM (No. Version 1.0)* [Online]. Imperial College Health Partners. Available: https://imperialcollegehealthpartners.com/wp-content/uploads/2018/07/Population-Health-Management-Flatpack-Version-1.0-Final-Sent.pdf [Accessed 1 May 2019].

NHS DIGITAL. Read Codes. https://digital.nhs.uk/services/terminology-and-classifications/read-codes

NICE. 2013. *How NICE measure value for money in relation to public health interventions* [Online]. Available: https://www.nice.org.uk/Media/Default/guidance/LGB10-Briefing-20150126.pdf [Accessed].

NICE. n.d.-a. *NICE guidelines* [Online]. National Institute for Health and Care Excellence. Available: https://www.nice.org.uk/About/What-we-do/Our-Programmes/NICE-guidance/NICE-guidelines [Accessed 1 June 2019].

NICE. n.d.-b. *Type 2 diabetes in adult overview* [Online]. National Institute for Health and Care Excellence. Available: https://pathways.nice.org.uk/pathways/type-2-diabetes-in-adults [Accessed 1 June 2019].

NUFFIELD TRUST. 2019. *Health & social care explained: NHS reform timeline* [Online]. Nuffield Trust. Available: https://www.nuffieldtrust.org.uk/health-and-social-care-explained/nhs-reform-timeline/ [Accessed 31 July 2019].

OSWALD, M., GRACE, J., URWIN, S. & BARNES, G. C. 2018. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'experimental' proportionality. *Information*

& Communications Technology Law, 27**,** 223-250.
https://doi.org/10.1080/13600834.2018.1458455

PRUSINSKI, S. 2017. Beyond risk scores. Prioritizing health care interventions with predictive analytics. *The Actuary.* The Society of Actuaries: USA.
https://theactuarymagazine.org/beyond-risk-scores/

ROLAND, M. & ABEL, G. 2012. Reducing emergency admissions: are we on the right track? *British Medical Journal,* 345**,** e6017. https://doi.org/10.1136/bmj.e6017

SCHRÖDER-BÄCK, P., DUNCAN, P., SHERLAW, W., BRALL, C. & CZABANOWSKA, K. 2014. Teaching seven principles for public health ethics: towards a curriculum for a short course on ethics in public health programmes. BMC Medical Ethics, 15, 73.https://doi.org/10.1186/1472-6939-15-73

SHEAFF, W., HALLIDAY, J., ØVRETVEIT, J., BYNG, R., EXWORTHY, M., PECKHAM, S. & ASTHANA, S. 2015. Integration and Continuity of Primary Care: Polyclinics and Alternatives, a Patient-Centred Analysis of How Organisation Constrains Care Coordination. Health Services and Delivery Research, 3, 1-150. http://dx.doi.org/10.3310/hsdr03350

SHEEHAN, M., HOPE, T., RHODES, R., BATTIN, M. & SILVERS, A. 2012. Allocating Health Care Resources in the UK. *Medicine and social justice: essays on the distribution of health care.* Oxford University Press.

SIKKA, R., MORATH, J. M. & LEAPE, L. 2015. The quadruple aim: Care, health, cost and meaning in work. *British Medical Journal BMJ Quality & Safety,* 24**,** 608-610.
https://doi.org/10.1136/bmjqs-2015-004160

SNOMED. http://www.snomed.org/

SNOOKS, H., BAILEY-JONES, K., BURGE-JONES, D., DALE, J., DAVIES, J., EVANS, B., FARR, A., FITZSIMMONS, D., HARRISON, J. & HEAVEN, M. 2018. Predictive risk stratification model: a randomised stepped-wedge trial in primary care (PRISMATIC).
https://doi.org/10.2196/jmir.5011

SPIKE, J. 2018. Principles for public health ethics. *Ethics, Medicine and Public Health,* 4**,** 13-20.

STATISTICS HOW TO. 2014. *nowball Sampling: Definition, Advantages and Disdvantage* [Online]. Available: https://www.statisticshowto.datasciencecentral.com/snowball-sampling/ [Accessed 21 October 2019].

STEVENTON, A. & BILLINGS, J. 2017. Preventing hospital readmissions: the importance of considering 'impactibility', not just predicted risk. 26(10)**,** 782-785.
https://doi.org/10.1136/bmjqs-2017-006629

TANNAHILL, A. 2008. Beyond evidence—to ethics: a decision-making framework for health promotion, public health and health improvement. *Health Promotion International,* 23**,** 380-390. https://doi.org/10.1093/heapro/dan032

UK BIOBANK LIMITED 2010. Biobank: Information Leaflet. *In:* LIMITED, U. B. (ed.). UK Biobank Limited.

UNITED KINGDOM GOVERNMENT. 2018a. *Guidance: Code of conduct for data-driven health and care technology* [Online]. London: Department of Health & Socail Care. Available: https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology [Accessed 01 December 2019].

UNITED KINGDOM GOVERNMENT. 2018b. *Guidance: Data Ethics Framework* [Online]. UK Government. Available: https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework [Accessed].

UNIVERSITY OF EDINBURGH. 2019. *Reflection Toolkit* [Online]. Edinburgh: University of Edinburgh. Available: https://www.ed.ac.uk/reflection/facilitators-toolkit [Accessed 1 May 2019].

VUIK, S. I. 2017. On the application of data-driven population segmentation to design patient-centred integrated care. https://doi.org/10.25560/58341

WAGNER, B. 2018. Ethics as an escape from regulation: From ethics-washing to ethics-shopping. *Being Profiled: Cogitas Ergo. Sum Amsterdam University Press, Amsterdam***,** 84-90.
https://www.jstor.org/stable/pdf/j.ctvhrd092.18.pdf

WHO. n.d.-a. *Constitution* [Online]. Geneva: The World Health Organization. Available: https://www.who.int/about/who-we-are/constitution [Accessed 1 May 2019].

WHO. n.d.-b. *Health Impact Assessment (HIA) - The determinants of health* [Online]. The World Health Organisation. Available: https://www.who.int/hia/evidence/doh/en/ [Accessed 9 July 2019].

WYATT, M. 2012. *ANGEL Taxonomy: a cognitive model for assessment, decision making and planning in complex care* [Online]. NHS Wales. Available:

www.metaph.org/sitesplus/documents/1065/ANGEL%20Taxonomy%20Briefing%20April%20
2012.pdf [Accessed 1 June 2019].

Appendices

Appendix 1: Interview Questionnaire used to identify practical considerations

General Approach

The questions we want to ask will depend on the knowledge of the expert. We envisage talking to at least four types of expert:

- Individual GP - at the coal face seeing the patients on a day to day basis
- "GP analyst" - analyses the data available to a GP (or GPs) to decide who is at risk
- Public Health analyst - analyses wider patient data, and uses (and sponsors) medical research, to provide guidance for practitioners and policy makers
- Specialist Consultant - at the coal face, applying specific expertise to treat individual patients, but also using expert knowledge to guide public health decisions

The language we use will be tailored to the prior knowledge of the individual with regard to risk segmentation and impactability

GP and GP Analyst

1. Do you have a systematic way to identify patients who are at risk of declining health?
2. How do you persuade patients to be aware of health risks/take medications/change lifestyle?
3. Does your approach vary depending on the patient?
4. If so, how does it vary, and what drives the variation?
5. Can you provide an example when your treatment decision varies depending on the likely adherence of your patient to different treatment options?
6. How do you judge the likely adherence of a patient to a treatment regime prior to applying it?
7. Do you assess whether patients are adhering to treatment regimes?
8. If so, how?
9. Are there commonly occurring situations in which the impact of a treatment on a particular patient is unpredictable - some patients respond very well, and others don't?
10. Are there commonly occurring situations in which the impact of a treatment on a particular patient is unmeasurable?
11. Are there any investigations you would like to see, that would help the medical fraternity understand which intervention will provide most benefit to a patient and how to apply the intervention in an optimal way for each patient?
12. Are there tools available to help you assess the potential impact of treatment on patients? Which ones do you use? Which ones don't you use? Why not?

Public Health Analyst

1. How do you measure the impact a current intervention is having?
2. How do you judge the impact an intervention will have before investing in it?
3. Are there examples where you have segmented the population by risk level to target those at highest risk?
4. Are there examples where you have segmented the population by likelihood of response to an intervention?
5. Do you use different types of intervention for different segments of the population?
6. Are there areas where you would like to make more use of segmentation by risk, but lack tools or other support? If so what do you lack?
7. Are there any interventions where further investigations should be done to assess the overall impact on a real population so that your decision making is better informed?
8. Do you need more resources to improve your data management or data analysis tools?
9. How do you implement a new intervention? What are the barriers and enablers to change?

Specialist Consultant (e.g. Oncologist)

1. How do you persuade patients to be aware of health risks/take medications/change lifestyle?
2. Does your approach vary depending on the patient?
3. If so, how does it vary, and what drives the variation?
4. Can you provide an example when your treatment decision varies depending on the likely adherence of your patient to different treatment options?

5. How do you judge the likely adherence of a patient to a treatment regime prior to applying it?
6. Do you assess whether patients are adhering to treatment regimes?
7. If so, how?
8. Are there commonly occurring situations in which the impact of a treatment on a particular patient is unpredictable - some patients respond very well, and others don't?
9. Are there commonly occurring situations in which the impact of a treatment on a particular patient is unmeasurable?
10. Are there any investigations you would like to see, that would help the medical fraternity understand which intervention will provide most benefit to a patient and how to apply the intervention in an optimal way for each patient?
11. Are there tools available to help you assess the potential impact of treatment on patients? Which ones do you use? Which ones don't you use? Why not?

Appendix 2: Impactability models

This appendix provides brief descriptions of the impactability models listed in Section 5. It includes the 4 models that are described and assessed in that section and also the models which, at this stage, have not been studied in further detail by the Working Party.

Traditional approaches

The traditional approaches include prioritizing patients with a high gap score, or weighted gap score. This means a gap in the care that the patient has received compared to recommended guidelines, such as those offered by NICE in the UK. Patients with the biggest gaps are prioritised, as they can be impacted by doing more of the things on the NICE guidelines.

One or more ambulatory care sensitive conditions - if the patient is managed optimally in community settings then they should not need hospital treatment for that condition. Examples of ACSC's are hypertension and asthma. Under this approach, patients with these conditions should be prioritised, as having an ACSC means they are impactable. This approach would require matching different ACSCs to different interventions.

Patients with certain stable characteristics could be viewed as having less need for intervention than patients whose corresponding conditions are more volatile. This is particularly true of interventions that are designed to stabilise a condition.

Extremely high- risk patients may be less amenable to certain interventions than lower risk patients; this might include many of the more preventative interventions. However, the opposite is likely to be true for other types of intervention that are designed to stabilise and manage very high-risk characteristics.

Finally, under the traditional approaches: identifying impactable moments. These are times at which the health or care professionals have contact with the patient for some other purpose but could optimize the contact time by offering wider services at these times.

Statistical approaches

"Statistical" approaches such as risk stratification and risk scores have been around in the health sector for some years. These models take advantage of the abundance of health data and computing power to run algorithms against big lists of patients. They aim to predict for individual patients the likelihood of an adverse event, such as requiring unplanned treatment e.g. non-elective hospital admissions (Lewis, 2015). These events are undesirable, costly, and sometimes preventable. Such models have been shown to be superior to other "case finding" approaches, including clinical judgement. Evidence suggests that "clinicians are less accurate than risk stratification tools at predicting risk" (Lewis, 2015).

Rising risk score takes the approach one step further by filtering on patients whose risk score has increased over a particular timeframe.

The next model type relates to comparing service utilization, risk characteristics and condition severity against benchmarks. This involves prioritising those patients whose condition is more severe than would be expected given their risk characteristics; or whose health system resource utilisation is higher than expected, for a given condition severity.

There are other advanced statistical methods to predict which high-risk patients will benefit most from any given preventive intervention. These include clustering techniques such as decision trees and neural networks. These methods can each be used with a number of programme designs.

Survey-based approaches

The third category covers alternative approaches such as conducting detailed questionnaires of individual patients. These approaches involve primary data collection i.e. the required data is not

routinely ready for extraction. Data collection implies a greater cost but this could be balanced with a relatively small patient segment. There are existing models which generate data at an individual level, including Patient Activation Measures (PAM) and ANGEL.

- PAM uses the level of knowledge, skills and confidence that a person has in managing their own health and care (NHS England, n.d.-a), as an indicator of impactability.
- The ANGEL score is a cognitive model for assessment, decision-making and planning in complex care (Wyatt, 2012), based on a qualitative assessment of the complexity of a patient's need across five dimensions: life and social activities, existing care needs, goals, escalation required to align needs and goals, and location. This could enhance understanding of the likely impact of an intervention on an individual's personal circumstances, while providing a framework for comparison across cohorts of people.

Appendix 3: Technical constructs supporting selected models

This appendix provides a mathematical framework for some of the models that were introduced in Section 5.

Impactability model 1: Prioritise patients with high treatment gap score

A mathematical model could be defined such that:

- i = 1 to n where n is the total number of treatments recommended in the NICE guidelines.
- Define RT_{ix} as recommended treatment i for patient x
- Define AT_i as actual treatment i for patient x
- Define G_{ix} as the gap score for recommended treatment i for patient x. $G_i = RT_i - AT_i$

A patient's total gap score would then be defined as $G_x = \sum_{i=1}^{n} G_{ix}$

Patients with a gap score over a pre-defined threshold would be considered 'impactable'.

Impactability model 2: Condition severity benchmarking models

Individuals or specific population segments can be benchmarked for a range of metrics covering risk characteristics of the individuals/segments, service utilisation and condition severity.

Method 2(a) involves identifying individuals/segments whose conditions are more severe than would be expected for their given risk characteristics.

Method 2(b) involves identifying individuals/segments who are using more resources (/whose service utilisation is higher) than would be expected for their given condition (/given their condition severity).

To illustrate the method, consider the following 3 population segments:

1. High direct measures of frailty despite low electronic Frailty Index (eFI) (Clegg et al., 2017).
2. Cardiovascular disease (CVD) events in people with low QRisk scores (Hippisley-Cox and Coupland, 2013).
3. Early development of complications of diabetes despite historically optimal control.

For a given condition/complexity/risk, there needs to be more than one possible treatment/intervention option to evaluate.

Impactability model 2(a)

This model identifies individuals or population segments with more severe conditions than would be expected for their risk characteristics, which may indicate unrecognised cofactors that merit intervention. In effect it indicates the presence of unidentified factors that contribute to risk and unrecognised gaps in care.

The following would be assessed for an individual:

- Measure risk, R_{ti}, at time, t, for individual, i
- Measure condition severity, C_{ti}, at time, t, for individual, i

The range for the condition severity would be considered:

If $c_1 < C_{ti} < c_2 \mid R_{ti} < R_{thresh}$ where R_{thresh} is a threshold for low-risk, then it would be concluded that the individual is "impactable" and should be offered the intervention.

The variables of c_1 and c_2 would be determined from data analysis and/or clinical judgement.

There could be several different interventions across the domain of C, relevant for each range. For example:

- If $c_1 < C_{ti} < c_2$ → then recommend intervention A
- If $c_2 < C_{ti} < c_3$ → then recommend intervention B
- If $c_3 < C_{ti} < c_4$ → then recommend intervention C

Outcome metric, O_{t+1}, at time t+1, for individual, i

If $o_1 < O_{t+1i} < o_2$, at time t+1 it would be concluded that the intervention was successful and the individual was, in fact, impactable.

The outcome metric could be the same as the condition severity metric at time t+1, i.e. $O_{t+1i} = C_{t+1i}$

Metric	High frailty	CVD risk	Diabetes
Risk R_{ti}, at time, t, for individual, i	Low eFI score at time t	Low QRisk score at time t.	Optimal historical measures of HbA1c, blood pressure, BMI, albuminuria.
Condition Severity C_{ti}, at time, t, for individual, i	Directly observed measures of frailty such as positive 'Timed up and go test', 'gait speed', 'grip strength', falls and hip fractures.	Cardiovascular events such as myocardial infarct, stroke, angina or revascularisation procedures.	Early development of complications of diabetes such as retinopathy, nephropathy, cataracts, or cardiovascular events.
Outcome metric O_{t+1}, at time t+1, for individual, i	Improvement in frailty events, or appropriate eFI reclassification.	Identification and treatment of unidentified co-morbidity	Identification and treatment of previously unidentified co-factors and reduction in rate of progression.
Description	Identification of people with directly observable measures of severe frailty or high utilisation despite having a low risk as calculated by eFI scores. This might prompt investigation for undiagnosed morbidity, exercise programmes, utilisation management and / or appropriate frailty classification intervention.	Development of premature cardiovascular disease despite low risk predicted by QRisk may be an indicator of unrecognised conditions such as autoimmune disease, diabetes, CKD, inflammatory disorders, sedentary lifestyle, poor diet, or alcohol abuse. This represents an opportunity to identify these problems earlier and intervene to prevent future events.	Identification of people with historically optimally controlled diabetes who nevertheless develop significant early complications of diabetes. This may be an indicator of unidentified co-factors for complications such as autoimmune disease, organic sleep disorders, and chronic pain (Kneale et al., 2018).

Table 8: Example metrics for models identifying individuals/segments whose conditions are more severe than would be expected for their given risk characteristics.

Risk is not the same as condition severity. Risk is also not considered to be the same as outcome. There may be circumstances where these items are defined to be the same thing e.g. eFI score can act as a measure of risk and as a measure of condition severity. However, this will not always be the case.

"Health risk" must be clearly defined and will need to include the following elements:
- Derived from characteristics of an individual person or segment;
- that can be measured
  - within a timeframe that allows preventative interventions to be made
- which are not in themselves an indicator of current bad health;
- but which have a strong correlation with future adverse health outcomes;
  - "adverse health outcomes" needs to be defined. E.g. unplanned hospital admission.

Model 2(a) - statistical components

Risk, $R_{ti} = F(a, b, c, d, ..., m)$

The risk metric is a function of predictor variables a, b, c, d, etc. which can be observed and measured for the individual patient. The function is to be fitted to the data by regression techniques, but it is a choice as to what type of function this is. E.g. a linear function, normal, etc.

If GLM theory is applied then need to choose:
- distribution function of output variable
- link function
- linear predictor.

Condition Severity, $C_{ti} = G(n, o, p, q, ..., z)$

The condition severity metric is a function of predictor variables n , o , p , q , etc. (which may or may not be the same as variables for the risk metric) which can be observed and measured for the individual patient. For the frailty example, it could be the eFI score.

Note that there are alternative statistical techniques besides GLM approaches, e.g. time-to-event models, or semi-/non-parametric models that relax some of the assumptions underlying GLM theory.

Impactability model 2(b)

This model would identify individuals or segments whose service utilisation is higher than would be expected given their condition severity. It is similar to method 2(a) but with a focus on healthcare utilisation as the 'severity' and the condition as the predictor of risk. Individuals with high utilisation despite low severity may be an indicator of unrecognised cofactors or comorbidity, or a need for utilisation management. Individuals with low utilisation despite high severity may be an indicator for unrecognised gaps in care.

Here the condition severity, C_{ti}, could be measured at time t for individual i.

Similarly, resource utilisation, U_{ti}, could be measured during time [t-1,t], for individual i.

This requires careful consideration of how to measure utilisation across different care settings. For example, it could be based on the costs of services (of GP consultations, Accident and Emergency Department visits, inpatient episodes) used by each patient. The issue of regression to the mean must be addressed to make this model successful.

If $u_1 < U_{ti} < u_2 | C_{ti} < C_{thresh}$ then it would be concluded that an individual is "impactable" and should be offered the intervention.

Both, u_1 and u_2 need to be determined from data analysis and/or clinical judgement.

There could be several different interventions across the domain of U, relevant for each range. For example:

- If $u_1 < U_{ti} < u_2$ → then recommend intervention X
- If $u_2 < U_{ti} < u_3$ → then recommend intervention Y
- If $u_3 < U_{ti} < u_4$ → then recommend intervention Z

For example, persons with high severity conditions and relatively low utilisation may be impactable by making services more accessible for these individuals. Conversely, persons with low severity conditions and relatively high utilisation may benefit from alternative care delivery such as community services instead of hospital-based services.

The outcome metric, O_{t+1} , is measured at time t+1, for individual, i.

If $o_1 < O_{t+1i} < o_2$, at time t+1 then it would be concluded that the intervention was successful and the individual was, in fact, impactable.

The outcome metric could be the same as the condition severity metric at time t+1, i.e. $O_{t+1i} = C_{t+1i}$

Metric	High frailty	CVD risk	Diabetes
Utilisation U_{ti}, during time [t-1, t], for individual, i	Number of GP visits, Outpatient visits, admissions, prescriptions, community nursing contacts, social care contacts, etc. The cost of these services, or a weighted measure of them.		
Condition Severity C_{ti}, at time t, for individual, i	Low eFI score at time t with high utilisation, or low eFI score and high utilisation.	Stable, well controlled CVD with high utilisation, or unstable, poorly controlled CVD with low utilisation.	Well controlled diabetes, but high utilisation rates, or poorly controlled diabetes and low utilisation rates.
Outcome metric O_{t+1} , at time t+1, for individual, i	Reduction in utilisation, appropriate recalibration of eFI score with associated intervention.	Reduction in utilisation, or stabilisation of CVD.	Reduction in utilisation, or improved control.
Description	The intervention could identify those whose utilisation is out of keeping with their condition severity. This may indicate an unrecognised gap in care with under-utilisation, or a need for utilisation management.		

Table 9: Example metrics for models identifying individuals/segments whose service utilisation is higher than would be expected for their given condition (severity)

These methods triangulate around condition severity (condition severity for a given risk level and service utilisation for a given condition severity). An alternative version would be to cut out condition severity and benchmark service utilisation for a given risk level directly.

Model 2(b) statistical components
Condition Severity, $C_{ti} = G\ (\ n\ ,\ o\ ,\ p\ ,\ q\ ,\ \dots\ ,\ z\ )$

I.e. The condition severity metric is a function of predictor variables n , o , p , q , etc. (which may or may not be same as variables for the risk metric) which can be observed and measured for the individual patient. For the frailty example, it could be the eFI score.

Method 3: Propensity to benefit scores
The propensity to benefit score is a method of prioritising high-risk individuals for treatment plans according to a score calculated for each individual taking into account the probability of engaging successfully with a programme, the potential to improve quality of care, and the potential to reduce costs. These latter considerations are the distinguishing factors versus mere risk stratification methods. The scores are calculated using a predictive model that has been calibrated to health data

using regression analysis. In one particular study, the approach was applied to insured members of a health plan who qualified for a high-risk case management programme (HRCM) (Hawkins et al., 2015). For any health condition with a specific preventive programme, a preventive programme may either benefit or be detrimental to the health of an individual. The benefit is a condition-specific measure and can be defined as a percentage. For example, for a patient with high cholesterol, the benefit of the preventive programme could be defined as the percentage improvement in their cholesterol level. For a patient with diabetes, the measured benefit of the preventive programme could be defined as the percentage improvement in their blood glucose level.

GLM (Generalised Linear Model) techniques can be applied to predict the benefit of a preventative programme for an individual before they join the programme. The main output of the predictive model is the "propensity to benefit score", or "propensity-to-succeed (PTS) score" (Hawkins et al., 2015). This latter method involved fitting logistic regression models to identify three groups of individuals: those most likely to engage in the preventative programme; those receiving high quality of care once engaged; and those incurring enough monetary savings related to programme participation to more than offset programme costs" (Hawkins et al., 2015). The independent variables in the regression analysis were membership details alongside medical and pharmacy claims data.

Since many different preventive programmes can exist for a specific condition, let's assume that there are 1:k preventive programmes for a specific condition. We can also introduce an index for a condition assuming there exist m conditions of interest.

Separate regression analyses would be run to derive bespoke ("PTS" or "PTB") scoring models for each preventative programme. The scoring model would then be applied to each individual who is being considered to be invited to participate in each programme. In the literature, the input variables were claims data as this method was assessed in an insurance-based health economy. A similar approach can be envisaged for non-insurance based health systems, i.e. where claims data does not exist but other types of data are available. It would require demographic, diagnosis and health service utilisation data about individuals as well as data related to the specific preventive programme. Data specific to the preventative programme could include details of phone calls planned and answered, in-home consultations, the individual's adherence/compliance to their prescribed medication, and whether or not the individual had a companion to accompany them to their medical appointment.

This model could be defined with the following variables:

- R_{tij} - the probability of an individual i at time t being at risk of developing a condition j
- C_{tij}, - measure the severity of condition j at time, t, for individual, i
- P_{tijk} - the probability that an individual i at time t with a condition j is likely to have improvements in levels of condition specific measures (i.e. benefit) from the preventive programme k
- Q_{tijk} - the probability that an individual i at time t with a condition j is likely to not to have improvements in levels of condition specific measures (i.e not benefit) from the preventive programme k such that $Q_{tijk} = 1 - P_{tijk}$
- S_{tijm} - the probability of an individual i at time t receiving treatment m for a condition j

Metric	Types of preventive programme	High frailty	CVD risk	Diabetes
Risk R_{tij}, at time, t, for individual, i with condition j		Age; lives alone or not; with/without social support.	Age, weight/BMI, blood pressure, smoker status.	Age, weight/BMI, blood pressure, smoker status
Condition Severity C_{tij}, at time, t, for individual, i, with condition j		eFI score at time t		HbA1c level at time t
P_{tijk} **propensity to benefit** for an individual i at time t with condition j from preventive programme k	Controlled preventive programme	Propensity of benefit of Increased physical activity (PA)	Changing the diet from unhealthy to healthy, losing weight, quitting smoking.	Changing the diet from unhealthy to heathy diet, losing weight
	Uncontrolled preventive programmes (Family history)		Prescription of cholesterol lowering medications such as statins.	Prescription of insulin or oral hypoglycaemic agents.

Table 10 Example metrics for models identifying individuals/segments with high propensity to benefit from intervention

Appendix 4: Details of the core ethical principles and list of the studies included in the ethics frameworks scoping review

Four core Principle of Population Health Ethics:

Beneficence and non-maleficence ('Do good' and 'do no harm')

Beneficence and non-maleficence are the key principles in traditional medical ethics and are often ascribed to the Hippocratic Oath (Spike, 2018).

In impactability modelling, benefits at the individual level are likely to have been ascertained and form inputs to the models, with or without the uncertainties associated with these quantities. Overall and group-specific benefits to the population tend to be explicitly modelled in impactability models and are expected to materialise if modelled interventions are implemented in the target population and setting.

In terms of non-maleficence, (do no harm) patient safety (including broader health and wellbeing outcomes) may be at risk of being compromised if there is a large degree of uncertainty in the model. Patient safety in specific population subgroups may be less well studied in non-pharmacological interventions, compared to pharmacological interventions such as statin use in diverse groups of people. Subgroup-specific evidence on safety of interventions based on impactability modelling should be sought. As well as the effects on patient safety from the interventions, potential ethical harms resulting from the whole PHM or impactability approach, such as compromising patient privacy and stigmatisation, should be considered.

Health Inequality and Fairness

Health inequality is perceived in many ways by the general public, for example through multiple ecological and individual cultural models (Elwell-Sutton et al., 2019), and across the wider determinants of health (WHO, n.d., Marmot et al., 2010, Court, 1981). It has been measured in diverse ways in sociological and social epidemiological studies, with the choice of measure normally based on data and contextual considerations. The Working Party has gathered its own perceptions of this topic through a pilot self-reflective exercise (Appendix 5). Additionally, laws on discrimination and data protection (Section 6.1) may be relevant to risk stratification.

With respect to health resource allocation, there are limited resources for health improvement and a subset of impactable groups may have to be selected for implementation, or a reduced level of the implementation selected. Resource allocation debates often represent competing views of justice, based on Dworkin's concepts of "brute luck" and "option luck":
1. Solidarity or "brute luck": mitigating misfortune through the shared risk among an entire population
2. Actuarial Fairness or "option luck": risk stratifying such that each person is only required to share the burdens of risk mitigation with others who have similar risk profiles (May et al., 2017)

Actuarial fairness focuses on risks that can be mitigated by conscious choices (May et al., 2017). It should be noted that health behaviours are not wholly conscious choices, as they are affected by policy and by social and environmental determinants of health (Bartley, 2016).

More practically, NHS England provides resources for organisations and the public to reduce inequalities (NHS England, n.d.-b). The Centre for Health Economics provides detailed guidance on the incorporation of measures of health inequalities in resource allocation models (Love-Koh et al., 2019). In decision making, the choice of intervention or policy alone may introduce inequalities, for example focusing on a health condition that disproportionately affects the well off. The prevailing use of QALYs in decision making in healthcare may disadvantage those who are already less advantaged, such as older individuals and those in poor health. The risk stratification methods proposed in the case studies (Section 5) are mostly based on individuals' health status, health behaviours or healthcare utilisation rather than socioeconomic, demographic or geographical characteristics. The stratification procedure itself may be affected by the underreporting of health conditions by less advantaged, and this phenomenon is difficult to measure. As individuals' health

status or behaviours can be a consequence of these types of inequality, care should be taken such that these inequalities are not exacerbated, and that the risk stratification does not use criteria that signify prior disadvantage (Braithwaite et al., 2016). For example, if patients are determined to be more impactable based on a specific level of healthcare utilisation, adjustments for patients' differential access to healthcare due to inequalities should be considered.

Justice in relation to procedures or the deliberative process is discussed separately in Section 6.

Independence/autonomy

The main area where patient autonomy is important, informed patient consent (Spike, 2018), is not as relevant in PHM as it is in clinical care and health data collection but is relevant where consent is required or recommended for the intervention in scope. Potential interventions based on patient activation measures should take into account that patients are informed of the modelling procedure and the consequences of the choices they make regarding their own health.

Although scientific autonomy does not feature as strongly as patient autonomy in biomedical ethics, it is highly important in PHM, where evidence-based decision-making is fundamental. Boden and McKendrick (2017) specifically stressed the importance of scientific independence in model development.

Other (more operational) ethical principles

Other principles that are not part of the four core population health ethics principles, but important to PHM and resource allocation, have been proposed. Schröder-Bäck et al. (2014) proposed three other principles: health maximisation (a focus on getting the most benefit for the population), efficiency (getting the most out of limited resources), and proportionality (restrictions on autonomy, freedom, or privacy are only justified if balanced by a greater gain in health benefits to society). The first two combine to represent value for money or utility, which relate to health economic considerations and PHM's aim to maximise value and minimise costs (May et al., 2017). Proportionality links the consideration of the four core principles together.

Transparency is discussed in NHS and government data ethics codes of conduct (Section 6.1) and as part of the principle of procedural justice (Spike, 2018), while openness and accountability form two of Tannahill's (2008) ten ethical principles for health promotion, public health and health improvement.

Procedural fairness or justice is highly relevant to both decision-making and modelling, especially when there are multiple modelling stages and involve stakeholders with established processes. This principle can be upheld by encouraging participation and defending the rights of all stakeholders in the decision making process, and giving processes the authority to stop the implementation of policies if there is enough opposition (Spike, 2018). It is also recommended that processes should be reasonable and accountable to the public, involve people with the relevant expertise and experience in decision making, and ensure decisions are made on the basis of proper understanding of scientific and clinical evidence (Dunn and Hope, 2018). Processes relating to procedural fairness are already in place in NICE (Dunn and Hope, 2018; Sheehan and Hope, 2012) and could be used as a starting point for impactability modelling. Dunn and Hope (2018) argue that the quality of the process, and not the judgement of medical ethicists or any other single member of the committee, ultimately determines how these principles should be reasoned through and applied in practice.

Studies included in the literature review of ethics frameworks:

1. BODEN, L. A. & MCKENDRICK, I. J. 2017. Model-based policymaking: A framework to promote ethical "Good Practice" in mathematical modeling for public health policymaking. Frontiers in Public Health, 5, 68. https://doi.org/10.3389/fpubh.2017.00068

2. COGGON, J. & VIENS, A. M. 2017. Public health ethics in practice: A background paper on public health ethics for the UK PHSKF [Online]. Public Health England. Available: https://www.gov.uk/government/publications/public-health-ethics-in-practice [Accessed 1 June 2019].

3. FADEN, R. R., KASS, N. E., GOODMAN, S. N., PRONOVOST, P., TUNIS, S. & BEAUCHAMP, T. L. 2013. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. Hastings Center Report, 43, S16-S27. https://doi.org/10.1002/hast.134

4. FOURIE, C., BILLER-ANDORNO, N. & WILD, V. 2014. Systematically evaluating the impact of diagnosis-related groups (DRGs) on health care delivery: a matrix of ethical implications. Health Policy, 115, 157-164. https://doi.org/10.1016/j.healthpol.2013.11.014

5. MAY, T., BYONANEBYE, J. & MEURER, J. 2017. The ethics of population health management: Collapsing the traditional boundary between patient care and public health. Population Health Management, 20, 167-169. https://doi.org/10.1089/pop.2016.0058

6. SPIKE, J. 2018. Principles for public health ethics. Ethics, Medicine and Public Health, 4, 13-20.

7. TANNAHILL, A. 2008. Beyond evidence—to ethics: a decision-making framework for health promotion, public health and health improvement. Health Promotion International, 23, 380-390. https://doi.org/10.1093/heapro/dan032.

Appendix 5: Template for a reflexive exercise

Exercise process

1. The exercise can take 1 hour to complete, with individuals encouraged to write answers down to encourage self-reflection. Answers are not shared or seen by anyone other than the individual. Being open, honest and transparent when answering is beneficial. There are no right or wrong answers. The exercise is designed to provoke thought, i.e. reflection. Times were allocated next to each question as a guide for the participant, but it is up to the individual how much time to spend considering each question.

2. Following 1, a group discussion is held. This is an opportunity to discuss the exercise and the experience as a group. Group-reflection was undertaken to increase the working knowledge in the group and to gain an appreciation and learn from others reflection experiences. It is important to establish with the group that not knowing or having answer to the self-reflection questions in 1 is acceptable; the learning is taken from the process of reflecting on the questions.

3. After the group exercise, major themes from the discussion can be circulated to the group. Areas not captured in the group discussion that members believe important can be facilitated through an anonymous survey platform. Actions can be agreed within the group following the discussion to further increase or improve understanding and awareness. These actions could be set at an individual, sub-group, or group level in a modelling or project team.

Sample questions

Questions for the self-reflexive exercises will need to be drafted. The nature of the items below will depend on the nature of the work and the group. Sample questions have been provided below but should not restrict the user.

1. Background
    a. The demographic profile of the group was reflected upon using an anonymous survey. This helped to understand the working group better. Primary data collection will require due consideration of GDPR.
    b. A survey of the ethics resources available to members was created to understand how comfortable members were with making ethics-based decisions and if support resources were available to the group.

2. Health
    a. Questions regarding health were posed to individuals to consider and for the group to reflect on.
    b. Some example questions include:

> - How would you define health?
> - What is a strength of the WHO definition for health (WHO, n.d.-a)? What is a weakness? How does your definition compare?
> - What does a holistic view of health mean to you?
> - What are some determinants of health? For you, your family, your colleagues and patients?

3. Impactability
    a. Questions regarding impactability were posed to individuals to consider and for the group to reflect on.
    b. Some example questions include:

- What does an impact to your health look like to you?
- What does '*creating an impact to health*' mean to you as a professional?
- What does '*impacting the populations' health*' look like?
- What do you believe currently creates an impact on population health? Positively and negatively?

---

4. Ethic scenarios
    a. Three discussion scenarios were generated for individuals to consider and for the group to reflect on.
    b. An example scenario to discuss is provided below:

Smoking has been banned in public places as second-hand smoke is detrimental to health. Taxes are imposed on tobacco with the aim of reducing consumption. Similarly, taxes are imposed on alcohol to reduce consumption.

- How is Minimum Unit Pricing of alcohol similar or different to tobacco?
- Reflecting on tobacco and alcohol, how does the Sugar Sweetened Beverage tax impact health inequalities?
- Reflecting on health interventions at a population level, does mandatory child vaccination raise ethical issues?

---

5. Stakeholders and perspectives
    a. Individuals were asked to consider different stakeholders and perspectives. These were then reflected on in the group setting.

Group-reflexive exercise

A facilitator will be required for the group exercise. This could be a member of the group or an independent person with qualitative research experience such as focus groups.

This is the setting under which the working party sought to establish actions resulting from the reflective exercises.

1. Establishing ground rules for group exercise
    a. There are no right or wrong answers. Full respect for peer input, values and comments. Never feel forced to share anything you are uncomfortable with and please be respectful of others' boundaries when sharing. No comments will be attributed to any person and to support open discussions Chatham House Rules are in effect.
    b. "It's only reflection if it strives toward a better understanding" (University of Edinburgh, 2019). The questions posed to the group are to help reflection on why others may have different perspectives of the same experience of the self-reflective exercise.
    c. The objective is to find purposeful examination of thoughts and practice.

2. Group reflection on background
    a. Questions were posed to the group regarding the outcome of this section. These included reflection and action orientated prompts.

Reflection prompts:
- What is your reflection on…?
- Do we perceive an issue with…?

Action orientated prompts:
- What action result from this group discussion?

---

3. Group reflection on health section and the impactability section in turn
    a. Questions were posed to the group regarding the outcome of this section. These included reflection and action orientated prompts.

Reflection prompts:
- What were some of the thoughts that you had that surprised you in this section?
- Were the questions easy to answer? If so, why? If not, why?
- Did you rely on personal or professional experiences to form your answers?

Action orientated prompts:
- How would we ensure that multiple stakeholders' views and perspectives are considered in our work?
- How would we as a group ensure we have a rounded view of …? Encompassing all the elements discussed?
- What action results from this group discussion?
- How would we as a group expand our understanding further?
- What would a good outcome of the work look like to you?

---

4. Group discussions were had on ethics scenarios
    a. Questions were posed to the group regarding the outcome of this section. These included reflection and action orientated prompts.

Reflection prompts:
- Did you feel there were issues in the scenario?
- Do you think the ethical issue could be navigated? How?
- What would improve the situation? What would make the situation worse?
- Why is it ok to…? Why is it not ok to…?

Action orientated prompts:
- How would we ensure that ethics is considered within our work?
- How will we ensure that health inequalities are considered within out work?

---

5. Actions
    a. The working party was posed questions on how their reflections would be featured in their work

www.actuaries.org.uk