Measuring General Practice Productivity – Detailed Project Description

Contents

Scientific abstract

There is a continuing need in the NHS to make the best use of resources available for the best possible outcomes. Despite most initial contacts happening within primary care, systems for measuring the performance of general practices are extremely limited, with the main mechanism for doing so (the quality and outcomes framework - QOF) criticised for being too narrow, with other indicators (ONS, NICE, CQC) similarly incomplete.

Productivity in health care is defined as a ratio of outputs to inputs, adjusted for the quality of outputs based on measures such as patient satisfaction, waiting times etc. Despite considerable development of measures of productivity covering secondary care, measures for primary care remain at best crude or at worst absent. In this study we aim to develop and test a measure of primary care productivity that can be applied across all typical general practices (including their multidisciplinary teams) in England. We will do this using a procedure based on the well-tested ProMES methodology (Productivity Measurement and Enhancement System), in which the views and experience of general practice staff and patients will help determine the form of the measure, which will then be tested for its feasibility. This will result in a robust and valid measure, developed by practitioners and public that can be applied across GP practices more widely.

The first stage will include ten workshops with practitioners and patients in 3 phases. Two large workshops in phase 1 will be used to explore what the key dimensions of effectiveness, productivity and quality are in general practice. In phase 2, six small workshops will explore different ways in which each could be measured, using existing data or otherwise. In phase 3 two large full-day workshops will consider the validity and relative strengths and weaknesses of the different measures identified, and how to weight the different dimensions so that an overall effectiveness measure can be constructed; this will then be converted to a productivity measure by dividing it by the input (expenditure of the practice). The resulting measure will then be put to a panel of experts, including representatives of NHS, regulatory, professional and patient bodies, to test its initial acceptability, and make adjustments if appropriate.

We will then recruit 50 general practices to undertake a feasibility study of using the measure. Each practice will apply the measure using an online system over a minimum of six months, tracking its overall productivity on a monthly basis. Using a range of evaluation methods, including an online survey with practice managers, and telephone interviews with a range of practice managers, GPs, other practice staff and patient representatives, we will investigate the acceptability of the measure, the time and resources taken to gather and enter the data, and any difficulties encountered. Recommendations about use of the measure will be made based on this evaluation.

The study will be supervised by a steering committee including representatives from NHS England, CQC, NICE, ONS, GPs, social care, and four PPI members. Participants (including GPs, other staff, and patients) will be recruited via a number of CCGs and locality managers who have agreed to participate in the project (to date six CCGs from different parts of England have expressed their willingness to participate).

Plain English summary

Primary care is a hugely important part of the NHS. General practitioners (GPs) and other staff in their practices provide the first point of care for the majority of patients, and continuing community-based care for many. Despite this, existing mechanisms for measuring the performance (effectiveness and productivity) of general practices are relatively crude and under-developed. The principal method is the Quality and Outcomes Framework (QOF), which is based around a series of indicators for a limited set of conditions that are readily measured or recorded. In order to maximise the benefit to patients and the public, productivity needs to be high, but at present this cannot be measured accurately, and it is difficult for patients and the public to know what quality standards they can expect from their GP and primary care.

This study will aim to correct this by developing a measure of productivity that can be applied across most general practices. It will do this using the principle that the people who can best dictate what needs to be included in the measure are those who are closest to the matter at hand: GPs, other healthcare staff, patients and the public. The first stage of the project will consult with a wide variety of people from these groups using a series of workshops; these will be based around a method called the Productivity Measurement and Enhancement System (ProMES), which has been widely and successfully used in many sectors including healthcare. In these workshops a productivity measure will be constructed gradually by first identifying and agreeing the important dimensions of effectiveness, then exploring how these are best measured, and then by identifying the importance of each.

After the productivity measure is developed, it will be tested in a feasibility study of 50 general practices. For at least six months staff in these practices will enter the necessary data into an online system which will map their own productivity on a monthly basis, testing out the ease of use. An online survey of practice managers, and a range of interviews with practice managers, GPs, other practice staff and patient representatives will evaluate the acceptability of the measure (including whether it is a fair reflection of the practice's effectiveness), the time and resources taken to gather and enter the data, and any difficulties encountered. Recommendations about use of the measure will be made based on this evaluation.

As well as the measure itself, a range of outputs will arise from the study, including summary reports for the participating practices, a final report setting out the findings and implications for policy makers and practitioners, practical guidance on implementing the measure for GP practices, CCGs and others to support them to implement the use of the measure in their practice, and various events and briefings for external audiences. Patient representatives will be involved in producing these outputs to ensure they can reach a public audience as well.

Background

Productivity and Effectiveness in Health Care

It is always a challenge for health care providers and administrators to be able to produce as much as possible with the resources available. Within the NHS, the Wanless Report (Wanless, 2002) identified that the NHS would need to be able to do far more, with the same resources, simply to maintain the same level of service – and that was at a time of relative prosperity and growth in the NHS. In times of relative austerity and uncertainty, the necessity becomes even greater.

Of course, in order to know how productive a service or unit is, it is first important to understand what productivity means. A classical definition of productivity is simply the ratio of outputs to inputs, and is often defined in simple financial terms. However, in health care this definition is not sufficient – the simple measurement of financial outputs does not usually take account of the quality of care delivered. Productivity in healthcare should measure "how much health for the pound, not how many events for the pound" (Berwick, 2005). Therefore a definition often used within health is "the ratio of outputs to inputs, adjusted for quality" (Appleby et al., 2010). The nature of this adjustment is a matter of debate, however – it is not generally possible to assess the financial effects of quality directly, as to do this would require assessment of the services' marginal contributions to social welfare (Bojke et al., 2013), and identifying and isolating these contributions would be difficult if not impossible.

While attempts have been made to measure quality-adjusted outputs directly (e.g. Castelli et al. 2007; Dawson et al., 2005), these have tended to focus on secondary care, and do not generally account for the wide range of potential data, but instead preferring routinely collected outcome data. Quality can refer to a mixture of things, including health outcomes, safety, and patient experience. We argue that, particularly for primary care, the full extent of quality cannot be measured without taking into account the views of patients, for example (see e.g. Appleby et al., 2004).

We therefore address the issue of productivity in two stages. In the first stage, we create a weighted measure of outcomes, or effectiveness. This is wider than what may be traditionally thought of as "productivity", but will consist of a range of quality-based outputs to be defined and specified as part of this research. These will be added together using weights also derived during the study to produce a single, composite, output measure (but one which can also be analysed in its component parts). In the second stage, we divide this effectiveness by a measure of inputs (resources used to generate these outcomes, measured in terms of practice expenditure) to give a productivity index. Both output and input measures will be indexed at a base period and changes measured over time. The overall productivity measure will therefore be the ratio of the output index to the input index.

Productivity and Effectiveness in General Practice

Primary care is a critically important sector of the NHS. General practitioners (GPs) represent the medical heart of primary care, often being the first port of call for patients, and with the advent of Clinical Commissioning Groups (CCGs) in 2013, GPs play a significant role in commissioning health care from acute and other services.

Access is a key issue for patients, but it is increasingly under pressure. Improved productivity is required in order to maximise the available capacity of primary care services and maintain access for patients. Recent research has demonstrated links between effective primary care delivery with various outcomes, including mortality rates (Honeyford et al., 2013; Levene et al., 2010; 2012), emergency department attendance (Baker et al., 2011), and emergency hospital admissions (Bankart et al., 2011). Despite this, systems for measuring the performance of general practices are limited, and existing measures are often criticised.

The current primary method for assessing general practice outputs is the Quality and Outcomes Framework (QOF). This sets payments to practices based on their activity against a number of indicators across multiple domains. The precise content has not remained stable from year to year: for 2014/15, QOF includes 73 clinical indicators and 11 public health indicators. Notably, all of the "quality and productivity" indicators and the one "patient experience" indicator from previous years have been retired from QOF in 2014/15, suggesting that only clinical outcomes, rather than other areas of effectiveness and patient experience, are being prioritised. QOF has been criticised in many ways, including being arbitrary in its setting of targets, influenced by contractual negotiations, being subject to regular changes and creating tensions between patient-centred consulting and management (Chew-Graham et al., 2013; Doran et al., 2014). Other output measures, such as those used by the ONS and NICE, likewise do not cover all activity (Massey, 2012).

The importance of primary care quality is further indicated by the fact that the Care Quality Commission (CQC) has recently begun inspecting general practices. This inspection includes out-of-hours services, asking the key questions of whether they are safe, effective, caring, responsive, and well-led (Care Quality Commission, 2014). This brings together quality and safety, but does not directly address productivity, and leads to a broad-brush rating at one of four levels between "inadequate" and "outstanding". A senior CQC inspector will be a key advisor to our project, further demonstrating the need for this research at the current time.

We contend that any comprehensive measure of general practice productivity or effectiveness would need to consider the wide range of outcomes from primary care, including elements relating to public health and health improvement. Baker and England (in press) have presented a framework covering many of these outcomes: both final outcomes (including mortality, morbidity, disease episodes, quality of life, adverse incidents, equity, patient satisfaction, costs, and time of work/school), and intermediate outcomes (e.g. clinical outcomes such as immunisation/screening, health behaviours, resource utilisation, patient experience and practitioner-related outcomes such as work satisfaction). For each outcome (or type of outcome) there is both a degree of importance of the outcome to the overall perception of effectiveness, and a degree of influence over which the general practice can have over it. For example, mortality is a very important final outcome, but there is a relatively small extent to which primary care impacts it directly. On the other

hand, patient satisfaction with care is increasingly seen as an important outcome, and the delivery of primary care will certainly influence this in a far more immediate way.

In order to capture the range of outcomes, but also the differing importance of them, a model is needed that addresses both of these. The model which we will use is ProMES – The Productivity Measurement and Enhancement System.

ProMES

The Productivity Measurement and Enhancement System (ProMES) was initially developed in the 1980s as a way to enable teams or work units to identify the factors that contribute to their productivity (or effectiveness), and to track this productivity over time, with feedback acting creating the motivation to improve (Pritchard et al., 1988). It involves four stages:

1. The first stage is to develop *objectives* (called "products" in the original terminology) – things that the unit (in this case, the general practice) is expected to do or produce. These would normally be determined by a series of meetings between members of the unit. Typically between 3-6 objectives might be identified, although this can vary depending on the type of work the unit does (and it may be substantially more for general practices)

2. The second stage is to develop *indicators* of the objectives – a way of measuring how well the unit is doing on each particular objective. These are developed by the same personnel who identify the objectives, and involve thinking of ways of identifying the extent to which the unit was doing well on a particular objective – either using existing data, or collective new data. Each objective would have at least one indicator, but may have more than one.

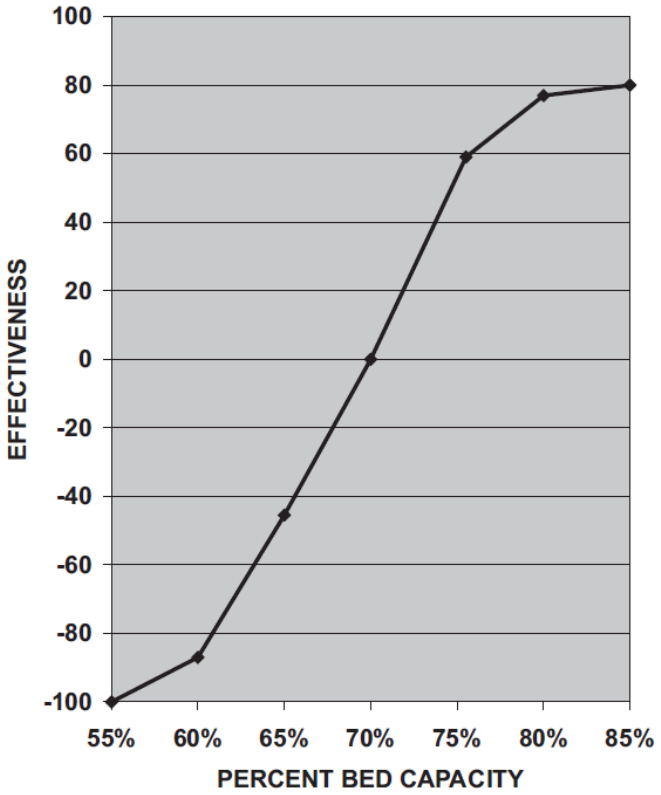
3. The third step is to identify *contingencies* – a method of weighting the different objectives. For each indicator, the contingency is a way of converting the actual value of the indicator into a score used for the overall productivity measure – in other words, saying just how good or bad particular values would be. These are set with a value of zero at the "neutral" point, with maximum and minimum values of up to +/-100 for the most important indicators, or proportionally less for less important indicators. These can be non-linear and asymmetrical (so that small changes can mean more at one point of the scale than at others). The setting of contingencies is done again as a collaborative effort between different unit members, although not necessarily the same ones as those who set the objectives and indicators. An example of a contingency is shown in Figure 1.

4. Finally, the system runs by collecting the indicator data over a designated period of time (e.g. a month), with an effectiveness score being calculated for each indicator at the end of that period, converted via the relevant contingency. These can then be summed to give an *overall effectiveness* score.

This comprises the *measurement* part of the process. The overall effectiveness score, as well as individual indicator effectiveness scores, are then fed back to unit members, leading to the *enhancement* part of the process – based on the theory of feedback and motivation (e.g.

Ilgen et al., 1979), the knowledge of not only the unit's overall effectiveness, but effectiveness on different objectives, creates a motivation to improve effectiveness. This has been shown to work with ProMES on many occasions: a meta-analysis of 83 field studies using ProMES found that there was a large and statistically significant improvement in performance following the beginning of ProMES feedback, with an average improvement of 1.16 standard deviations.

Figure 1: Example of a ProMES contingency (reproduced from Pritchard et al., 2008)[1]



The original ProMES methodology is designed to undertake this process with one team or work unit only. Obviously, with the objective of developing an effectiveness measure that can apply to multiple teams, this specific method would not work. Therefore we intend to use an adaptation of this method for multiple teams, in which representative members are brought together to undertake stages 1 to 3 of the ProMES approach as described above. Large scale adaptations have been used before in NHS settings (e.g. Poulton & West, 1993; Richards & Rees, 1998; West et al., 2012). The precise details of this adapted methodology will be described in our research design below.

[1] NB Although this example comes from a health care setting, it is not intended to suggest that percent bed capacity would be an appropriate indicator for general practices

Aims and objectives of this study

The main aim of this research is to develop and evaluate a measure of productivity (a ratio of quality-adjusted effectiveness to inputs) which can be applied across all typical general practices in England, and which may result in improvements in practice leading to better patient outcomes.

Specifically, the objectives are:

1. To develop, via a series of ProMES-based workshops with primary care providers and patients, a standardised, comprehensive measure of general practice productivity
2. To test the feasibility and acceptability of the measure by applying it in 50 general practices over a 6-month period
3. To use this evaluation to produce recommendations about the wider use of the measure across primary health care in consultation with key stakeholders at local and national level

Design

The study will take place over 30 months (with an intended start date of 1 June 2015). There are two main stages to the study, along with one preliminary phase:

1. Development of the measure
    a. ProMES-based workshops
    b. Finalisation of measure
2. Feasibility study of the measure
    a. Application of the measure in 50 general practices
    b. Evaluation of the application and reporting

Months 1-6: Preliminary phase

This preliminary phase will be used to complete all pre-fieldwork tasks, including:

- Completion of ethics and research governance approvals (these will begin before month 1, upon confirmation of successful award)
- Recruitment of participants for stage 1 workshops
- Finalising the design of stage 1 workshops

Months 7-15: Stage 1 – Development of the measure

The main objective of stage 1 is:

1. To develop, via a series of ProMES-based workshops with primary care providers and patients, a standardised, comprehensive measure of general practice productivity

Specifically, this breaks down into two sub-objectives:

i. To run a series of ten workshops, based on the ProMES methodology, involving primary care providers and patients, in order to generate

components of a measure of effectiveness and the comparative weighting of these components

ii. To produce a measure of effectiveness, and hence productivity, to be the subject of a consensus exercise involving a panel of key stakeholders, experts in primary care and patient representatives

iii. To create an online version of the measure using Ability© software with a view to it being applied within practices nationally.

Sub-objective 1i will be achieved via stage 1a; sub-objectives 1ii and 1iii via stage 1b.

Months 7-12: Stage 1a – ProMES-based workshops

As previously described in the background section, the ProMES (Productivity Measurement and Enhancement System) approach involves the identification of objectives, indicators of those objectives, and contingencies to convert indicators into meaningful effectiveness scores. Although in the original process only one team at a time would do this, it has been successfully adapted to larger scale settings, in which representatives from larger groups of teams work jointly to develop a common measure (e.g. West & Poulton, 1993; Rees & Richards, 1998; West et al., 2012). Two of the investigators from this proposal (Jeremy Dawson, Michael West) previously worked on these studies.

This stage, therefore, builds on experience of the previous research while remaining faithful to the original ProMES methods. It does this by using ten workshops in three phases:

- Phase 1: Two large full-day workshops, each involving a mixture of participants (around 40 each), including patients, GPs and other practice staff (both clinical and non-clinical), and other stakeholders (including representatives from national bodies). A series of exercises (some involving separate interest groups, some involving mixed smaller groups, and some involving all participants) will explore the key dimensions of effectiveness, identifying the objectives of a general practice team. These will then be analysed and consolidated, creating the ProMES objectives

- Phase 2: Six small half-day workshops, involving 4-8 participants each. In five of these, the participants will be primary care professionals, including GPs and other staff; in the other, they will be formed of PPI representatives. Each workshop will consider a subset of the objectives identified in phase 1, and identify potential indicators for each. These indicators may involve existing data (e.g. QOF indicators, HES data, General Practice Extraction Data, e.g. MIQUEST), or if necessary other data that could be collected.

- Phase 3: Two large full-day workshops, involving a similar group of participants to Phase 1. The first part of each workshop will consider the face validity and relative strengths and weaknesses of the indicators identified in phase 2. The remainder will be used to create the contingencies for the measure: used for converting and weighting the indicators so that an overall effectiveness score can be constructed. This will make use of on-screen display of contingencies (such as that shown in

Figure 1, above) to enable comparison and consolidation of differing opinions, using the Ability© software (described below).

The workshops will be facilitated by one of the investigators (Lee Adams) who has substantial previous experience of facilitating large-scale workshops, and supported by other investigators and project staff. The phase 3 workshops will be supported and co-facilitated by the makers of the Ability© software, which is designed to make the ProMES method easily achievable online, and is described fully in Stage 2 below.

The methods for the workshops will be similar to that used by West et al. (2012), but the design will be finalised in stage 1b.

Although this stage will mainly take place between months 7 and 12, recruitment of participants will of course begin before this. The workshops themselves will occur between months 7 and 11, with analysis occurring between each phase of workshops and final decision making about the measure (involving the steering group) in month 12.

Months 13-15: Stage 1b – Finalisation of measure

We will analyse the results from the workshops and integrating findings to create a single version of the measure, which will measure effectiveness: this is the numerator part of the productivity measure, i.e. the quality-adjusted output that will then be divided by the input. (Input will be measured as total spend per practice, and will be described more clearly in stage 2.)

We will then undertake a consensus exercise to determine whether this measure appears both feasible (in terms of the data collection required) and valid (i.e. that it appears to measure something that is accepted as a measure of productivity). This will take place in the form of an initial meeting, to which we will invite representatives from relevant NHS, regulatory, professional and patient bodies, including NHS England, Public Health England, the Care Quality Commission, the Royal College of General Practitioners, the British Medical Association, National Association of Primary Care, Department of Health, Local Government Association, Clinical Commissioning Groups, Healthwatch, Patients Associations, directors of adult social care services and of children's services., The measure will be presented and feedback sought about its feasibility and validity. The proposed sources of data in particular will be examined for their quality and availability: if the view is that proposed data will be of poor quality or may be difficult to extract, then this will either be monitored closely during the feasibility study (stage 2a) or replaced, depending on the level of concern. Following this, and any adjustments that may be needed, subsequent iterations of the measure may be presented to the same people via email in order to gain acceptability.

During this period we will also create an online portal in which practices will be able to apply the measure (enter the data for each indicator, and get a monthly report on practice productivity). We will do this using a tailored version of the software Ability©, which is designed for ProMES-style measures. Ability© is software developed by BlackBox Open (www.blackboxopen.com), a German company, to assist teams with the ProMES process. In its original version it presents teams with a series of questions designed to elicit objectives and

indicators (in our study this will be done using the first two phases of the ProMES workshops). It then moves to develop the contingencies for each indicator, using a series of structured questions and interactive manipulation of contingency graphs, such as that shown in Figure 1. This element of the software will be adapted to fit our phase 3 ProMES workshops, so that different groups can identify their own contingencies and these can then be compared.

We will also use this stage to finalise recruitment of the 50 practices we will need for stage 2, and prepare the documentation needed for this.

The main output of Stage 1 will be a measure (comprising individual indicators and a composite overall measure) that can be applied by individual general practices, using routinely collected or otherwise easily-collected data, with a piece of tailored software.

Months 15-30: Stage 2 – Feasibility study of the measure

The objectives for this stage are:

2. To test the feasibility and acceptability of the measure by applying it in 50 general practices over a 6-month period
3. To use this evaluation to produce recommendations about the wider use of the measure across primary health care in consultation with key stakeholders at local and national level

Specifically, objective 2 breaks down into the following sub-objectives:

i. To enable 50 general practices to use the measure over a six-month period
ii. To evaluate how the measure is perceived by the general practice community and patient representatives, including whether it is seen as a valid reflection of actual performance
iii. To determine how much time and other resource is needed to gather and enter the data within practices
iv. To determine whether any specific difficulties are faced by practices in gathering the data and using the measure
v. To evaluate whether practice staff found the feedback produced by the measure software useful
vi. To examine whether improvements in productivity were observed over the six-month period

The application of the measure in objective 2 will be done during stage 2a, with the evaluation during the (overlapping) stage 2b.

Months 16-24: Stage 2a – Application of the measure

This stage will be used to apply the measure across a pilot sample of 50 varied primary care teams. Before the start of this stage we will recruit 50 general practices to pilot the measure. These will be selected using a two-stage process. First, between six and eight Clinical Commissioning Groups (CCGs) will be invited to participate, covering different geographical areas of England (to cover a minimum of four different local government regions and a mixture of more urban and more rural areas). Within each we will recruit up to 10 practices; these will be purposively selected to ensure significant variety across at least the following metrics: (i) size of practice; (ii) deprivation of local area; (iii) demographic make-up of local area (including in particular variation in ethnicity and average age). We will also ensure that different models of practice ownership are included within the sample.

Each practice will then apply the measure over a minimum of a 6 month period (we expect the start of this period to be staggered between practices over around three months). This will be facilitated by using the tailored version of the Ability© software produced in stage 1b. The online software will provides a portal into which teams can, on a monthly basis, input the data from the different indicators, and the software will return an overall effectiveness score, as well as an effectiveness score for each objective and each indicator. Crucially, the indicators and contingencies will be set by the research team and advisory group following analysis of Stage 1, so our adapted version of the software will lock these centrally, and allow the research team to extract indicator data directly. However, it will also allow teams to set additional objectives, indicators and contingencies, which will not form part of the core effectiveness measure, but might be of specific interest to a particular team and so provide feedback on those. The use of this measure will be supported and monitored by a research associate.

We will offer each practice a payment of £500 for their participation in the study, in order to recompense the time of staff entering data into the software and performing other related activities. We will ensure we are in contact with each practice regularly during the six months so that we are aware of any problems that may be faced.

Months 22-30: Stage 2b – Evaluation of the application and reporting

There will be two main elements to the evaluation: a survey of practice managers, and telephone interviews with a range of practice managers, GPs, other practice staff and patient representatives.

Survey of practice managers

An online survey will be conducted at the end of the six month period. This will serve the joint purposes of addressing study objective 2 (primarily sub-objectives 2iii and 2iv, although 2ii and 2v will also be included), and of gathering data for the denominator of the productivity calculation – the inputs (total spend by the practice).

The questionnaire will ask about the amount of time spent on using the measure in the practice, any other resources that were needed to apply the measure, any obstacles in using the software (we will have some idea of any obstacles from the monitoring over the six-

month period), as well as including rating scales on the perceived usefulness of the measure, and the automatic feedback generated by the software.

We will also ask for the total expenditure per month by the practice during the period, enabling a productivity calculation. Practices will, under changes to the GMS contract in 2015, be obliged to publish annual accounts. We will ask about fluctuations in monthly expenditure, e.g. different numbers of staff employed at different times. If no fluctuations exist we will calculate expenditure as the annual expenditure divided by 12. If fluctuations do exist, then we will ask for further detail of this to enable a suitable denominator for the productivity measure to be provided.

Telephone interviews

We will conduct 50-60 telephone interviews to explore in more detail the perceived usefulness of the measure (whether the measure itself has face validity and is a fair reflection of the objectives or 'product' of their practice), any difficulties faced in using the measure/software, and the automatic feedback generated by the software.

Interviewees will include 15-20 practice managers, 10-15 GPs, 10-15 other practice staff, and 10-15 patient group representatives. We will attempt to ensure that all 50 practices are included in this evaluation; some practices will have more than one interviewee. Interviews with patient group representatives will focus on the perceived usefulness aspect of the evaluation.

In addition, we will study the values of the measure reported by each practice over the six-month period. According to a meta-analysis of ProMES methods (Pritchard et al., 2008), the feedback aspect implementation is likely to lead to increased effectiveness due to the motivation this provides. If we observe any changes of this nature, it may give rise to a subsequent trial or other study to examine whether such changes are genuinely due to the use of the ProMES methods.

The final part of stage 2 will involve the analysis and reporting on the results and making recommendations. This will be described in more detail in the section on dissemination and outputs.

Patient and Public Involvement

Given the importance and relevance of primary care to the public, it is essential that patient and public involvement (PPI) in this study is not only included, but is embedded throughout. Patient experience is anticipated to be one of the key outcome domains, and so ensuring that the patient and public view is included at all stages, from the design of the workshops through to the evaluation and reporting, is crucial. Indeed, one of the co-investigators (Amanda Forrest) is a PPI representative; she has had substantial input into the design of the study and will continue to do so throughout the study itself.

Besides a PPI investigator, the principal mechanisms through which there will there will be PPI input into the study are as follows. All patient and public involvement will be compensated in accordance with INVOLVE guidelines.

PPI representation on steering group

We will include four PPI representatives as part of our project steering group. These representatives will be recruited (via Healthwatch organisations) so that they represent a diversity of geographical locations, and as far as possible of other demographic characteristics as well. They will be recruited before the study start date, so that they will be involved from the very beginning of the study: this is important so that they have involvement in the design of the stage 1 workshops, as well as any relevant input into the ethics & governance applications. They will also help by advising over recruitment of patients for the latter stages of the study, as well as participating fully in advisory roles over all other elements of the project. Whether or not the PPI representatives have been involved in previous research, we will be very clear as to the scope of the study, and lay out the ground rules as to what elements of the study they (and the steering group as a whole) are able to influence, and which elements are fixed (contractually or otherwise), so that expectations are realistic on both sides.

Patient involvement in ProMES workshops

Patients will form a large proportion – up to 50% – of the participants in the ProMES workshops during stage 1 of the research. It is essential that the perspective of the patient experience is taken into account in determining the domains of effectiveness, as well as the potential measures and weighting of these measures. Therefore the 70-80 patients who are involved in these workshops will be recruited using a variety of means (including Healthwatch and patient groups associated with the general practices participating in the study) so that they include minority groups and hard-to-reach communities (e.g. disabled and BME patients). The phase 1 and phase 3 workshops will be structured so that some of the exercises are based in small groups of similar individuals – e.g. tables of patients – and some involve mixed groups (i.e. mixing patients with GPs, other practice staff etc.). This will ensure that the patient voice is heard clearly and is also taken into account by other groups.

Evaluation of application of measure

As part of the evaluation of the application of the measure, interviews will be conducted with patient group representatives, which will predominantly examine the perceived usefulness of the measure produced.

Reporting and dissemination

Amongst the various reports and other dissemination strategies (described below), we will ask relevant PPI representatives (who may or may not be members of the steering group) to participate in two specific ways:

- Writing articles for patient-focussed outlets (we have allowed funds for 3 days of writing time for this)
- Two patients speaking about the study at a relevant conference (or a speaker at two different conferences if that is thought to be a better use of funds)

In addition, all PPI representatives on the steering group will have the opportunity to review and comment on other project outputs, including (but not limited to) the full project report.

Dissemination and output

Several routes will be used to share findings, led by the King's Fund, which has strong relationships with NHS management and clinical community, policy makers and the media, and a proven track record of applying research findings to shape policy and practice, as well as extensive internal resources and technologies.

As well as dissemination to policy makers, regulators and practitioners in the wider health and social care system, we will also work with patient leaders and organisations including Healthwatch (locally and nationally), as well as local government (including the Local Government Association), and Health and Wellbeing Boards to disseminate findings and show what can be expected from GPs/primary care . Activities will include a range of web based material, events or webinars, articles in popular journals and social and popular media.

Interim findings will be fed back to participating practices through a short report and/or presentation, designed to draw out issues and recommendations about using the measure and will encourage ongoing learning during the testing phase. We will also produce a presentation and short practitioner-oriented report on emerging findings from the research for use by policymakers and NHS leaders (e.g. CQC, NHS England, Royal College of General Practitioners, National Association of Primary Care). This will focus on recommendations for national bodies to support implementation of productivity measures. We will also develop some case studies of participating sites outlining their experience of using the measure.

The main study report will draw out findings of practical use to practitioners and policy-makers. We will also develop an accessible online guide for GP practices, CCGs and others to support implementation of the measure. This will include the benefits that GPs may realise from adopting the measure, simple steps they can take to use it in their practice and where they can get further support.

We will undertake normal academic forms of dissemination though publications in peer-reviewed journals and presentations at academic forums. We will also publish in trade press to access a practitioner audience. Through existing relationships with CQC and Royal Colleges we will consult on developing the measure for national use to assure and promote continually improving, high quality primary care. We will consult with local GP practices and patient groups to ensure the measure is developed in a way that helps improve productivity and quality, well-being of staff, and patient experience of high quality and compassionate care.

Outputs will include:

- Brief summary reports from the workshop and engagement activity throughout the project lifespan

- Presentation/paper containing interim findings including learning to date for the participating practices
- Briefing and case studies on emerging findings for external audiences
- Final research report setting out the findings and implications for policy makers, practitioners and further research required
- Practical guidance on implementing the measure for GP practices, CCGs and others to support them to implement the use of the measure in their practice
- A series of communications to trade press, health policy and practice leaders, and disseminated to key audiences including relevant stakeholder groups.
- Speaking at events following publication of the report to promote the messages to key policy makers
- Patient-led presentations and blogs to ensure our findings are available to a more general audience

Our approach aims to build up interest and debate about measuring productivity in primary care from the start of the project and well beyond the project end, to garner the maximum impact. We recognise that we can only initiate debate, discussion and activity, and that others will need to want to support this work for this to have greater reach and impact. For this reason, we will develop a stakeholder engagement strategy which will identify potential audiences (such as primary care practitioners, CCGs, regulators and policy makers and the patient community) and specific targeted activity we would undertake to enthuse and involve them in the project. This will include a range of online, face-to-face, written and event based activity and utilising the patient involvement groups in co-design and dissemination.

In particular we will ensure NHS England, CQC, Healthwatch and Public Health England are fully briefed of the findings and involved in our research, to inform their regulatory and performance monitoring functions. This builds on work we are already doing with CQC to inform their new regulatory approach to primary care inspection and assessment. Through this work, we will work with policy makers to explore the potential for any measure that is developed to become a standard part of the regulatory and assessment process for primary care. Developing practical guidance for primary care teams on using the measure that also sets out clear benefits for practitioners may also further embed the measure into primary care practice.

Management of project

The project will be led by Jeremy Dawson (University of Sheffield), who will take day-to-day responsibility for the successful delivery of the project. He has worked closely with the investigators at the other institutions previously, most notably Michael West (The King's Fund) with whom he has worked for 15 years on numerous projects, and who will lead The King's Fund section of the work, and also provide mentorship to JD in his role as Principal Investigator.

The study team will meet face-to-face (in Sheffield or London) at six-monthly intervals; however, there will be monthly teleconferences in between to ensure regular communication of information.

The study will be supported by a steering committee, composed in accordance with NIHR guidelines, which will meet on three occasions during the project (in the first three months, before stage 2 to approve the finalisation of the measure to be tested, and towards the end of the study). This steering group will include representatives from NHS England, CQC, NICE, the Office for National Statistics, GP bodies, and four PPI members, as well as an expert in ProMES.

Project team member expertise

Jeremy Dawson will lead the project and will manage the research associate at Sheffield. His previous experience includes leading or participating in several projects which involve measurement in health care organisations. He is a statistician by background.

Michael West has led many large-scale research projects for NIHR and others. He has worked with GP practices in three previous projects. He will lead the work done at TZhe King's Fund, will provide psychological input for the ProMES workshops, and will provide mentorship for Jeremy Dawson as principal investigator.

Richard Baker is a GP and Professor of Quality in Health Care, and his input will ensure the study team has expertise in general practice from the medical side, as well as an eye on care quality. **John Appleby** is chief economist at the King's Fund and will provide health economics expertise when developing the productivity measurement.

Lee Adams has a background as a senior manager in local government and has 30 years' experience in the NHS as a senior commissioner and Public Health Specialist. She will lead the facilitation of the ProMES workshops. **Amanda Forrest** is a lay member of Sheffield CCG, has worked in various other organisations with patient and public involvement, and will lead the PPI input for the project. **Paul Wike** and **Michelle Wilde** are locality and practice managers in Sheffield, will provide input from the perspective of practice managers, and will assist with recruitment of participants for all phases.

Timeline

Stage	2015								2016												2017											
	Pre	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	Post
Preliminary																																
Ethics & governance																																
Finalising workshop design																																
1a: ProMES workshops																																
Recruitment of participants																																
10 workshops																																
Analysis																																
1b: Finalisation of measure																																
Consensus workshop																																
Finalisation of measure																																
2a: Application of measure																																
Recruitment of practices																																
Implementation																																
2b: Evaluation of measure and reporting																																
Practice manager survey																																
Telephone interviews																																
Analysis																																
Interim reports																																
Final report writing																																
Further dissemination																																
Research management																																
Finalise steering committee																																
Steering committee meeting																																
Project team meetings																																

References

Appleby, J., & Devlin, N. (2004). Measuring success in the NHS: Using patient-assessed health outcomes to manage the performance of healthcare providers. London: The King's Fund.

Appleby, J., Ham, C., Imison, C., & Jennings, M. (2010). *Improving NHS Productivity: More with the same not more of the same*. London: The King's Fund.

Baker, R., & England, J. (in press). Should we use outcomes data to help manage general practice? *British Journal of General Practice.*

Baker, R., Bankart, M. J., Rashid, A., Banerjee, J., Conroy, S., Habiba, M., & Camosso-Stefinovic, J. (2011). Characteristics of general practices associated with emergency-department attendance rates: a cross-sectional study. *BMJ Quality & Safety*, *20*(11), 953-958.

Bankart, M. J. G., Baker, R., Rashid, A., Habiba, M., Banerjee, J., Hsu, R., & Wilson, A. (2011). Characteristics of general practices associated with emergency admission rates to hospital: a cross-sectional study. *Emergency Medicine Journal*, *28*(7), 558-563.

Berwick, D. M. (2005). Measuring NHS productivity: how much health for the pound, not how many events for the pound. *BMJ: British Medical Journal*, *330*(7498), 975.

Bojke, C., Castelli, A. et al (2012*). Productivity of the English National Health Service 2003-04 to 2009-10*. Centre for Health Economics, CHE Research Paper 76. Retrieved from http://www.york.ac.uk/che/publications/in-house/, 10 September 2014.

Care Quality Commission (2014). *A fresh start for the regulation and inspection of GP practices and GP out-of-hours services*. Retrieved from http://www.cqc.org.uk/about-us/services-we-regulate/doctorsgps, 15 April 2014.

Castelli, A., Dawson, D., Gravelle, H., Jacobs, R., Kind, P., Loveridge, P., & Weale, M. (2007). A new approach to measuring health system output and productivity. *National Institute Economic Review*, *200*(1), 105-117.

Chew-Graham, C. A., Hunter, C., Langer, S., Stenhoff, A., Drinkwater, J., Guthrie, E. A., & Salmon, P. (2013). How QOF is shaping primary care review consultations: a longitudinal qualitative study. *BMC Family Practice, 14(1),* 103.

Dawson, D., Gravelle, H., O'Mahony, M., Street, A., Weale, M. et al. (2005). *Developing new approaches to measuring NHS outputs and productivity*. National Institute for Economic and Social Research, London & Centre for Health Economics, University of York, Research Paper 6. Retrieved from http://www.york.ac.uk/media/che/documents/papers/researchpapers/rp6_Measuring_NHS_outputs_and_productivity.pdf, 10 September 2014.

Doran, T., Kontopantelis, E., Reeves, D., Sutton, M., & Ryan, A. M. (2014). Setting performance targets in pay for performance programmes: what can we learn from QOF? *British Medical Journal, 348,* g1595.

Honeyford, K., Baker, R., Bankart, M. J. G., & Jones, D. (2013). Modelling factors in primary care quality improvement: a cross-sectional study of premature CHD mortality. *BMJ Open*, *3*(10), e003391.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of applied psychology*, *64*(4), 349.

Levene, L. S., Baker, R., Bankart, M. J. G., & Khunti, K. (2010). Association of features of primary health care with coronary heart disease mortality. *JAMA*, *304*(18), 2028-2034.

Levene, L. S., Bankart, J., Khunti, K., & Baker, R. (2012). Association of primary care characteristics with variations in mortality rates in England: an observational study. *PloS One*, *7*(10), e47800.

Massey, F. (2012). *Public Service Productivity Estimates: Healthcare, 2010*. London: Office for National Statistics.

Poulton, B. C. & West, M. A. (1993). Primary health care team effectiveness: developing a constituency approach. *Health & Social Care, 2,* 77-84.

Pritchard, R. D., Harrell, M. M., DiazGranados, D., & Guzman, M. J. (2008). The productivity measurement and enhancement system: a meta-analysis. *Journal of Applied Psychology, 93(3),* 540.

Pritchard, R. D., Jones, S. D., Roth, P. L., Stuebing, K. K., & Ekeberg, S. E. (1988). Effects of group feedback, goal setting, and incentives on organizational productivity. *Journal of Applied Psychology, 73(2),* 337.

Richards, A., & Rees, A. (1998). Developing criteria to measure the effectiveness of community mental health teams. *Mental Health Care, 2(1),* 14-17.

Wanless, D. (2002). *Securing our Future Health: Taking a long-term view. A final report*. London: HM Treasury.

West, M. A., Alimo-Metcalfe, B., Dawson, J. F., El Ansari, W., Glasby, J., Hardy, G., et al. (2012). *Effectiveness of Multi-Professional Team Working (MPTW) in Mental Health Care. Final report.* NIHR Service Delivery and Organisation programme.