Data Quality Evaluation for the Summary Care Record

An independent evaluation by University College London

Emma Byrne, Dipak Kalra, Katja Stramer, Trisha Greenhalgh

28 May 2008
Contents

| ACKNOWLEDGEMENTS | 3 |
| PREFACE | 4 |
| 1. EXECUTIVE SUMMARY | 5 |
| 1.1. SCOPE AND CONTEXT | 5 |
| 1.2. KEY FINDINGS | 5 |
| 1.3. KEY RECOMMENDATIONS | 7 |
| 2. INTRODUCTION | 9 |
| 2.1. REMIT | 9 |
| 2.2. BACKGROUND TO DATA QUALITY CONCERNS IN RELATION TO THE SCR | 10 |
| 2.3. DEFINING DATA QUALITY STANDARDS | 11 |
| 2.4. CONTENT OF THE REPORT | 13 |
| 2.5. METHODS | 13 |
| 3. USE CASE SCENARIOS FOR THE SCR | 16 |
| 3.1. THE RANGE OF USE CASES | 16 |
| 3.2. USE CASES IN THE CLINICAL SETTING | 16 |
| 3.3. DELPHI DATA CONCERNING USE CASE PRIORITIES | 19 |
| 4. LITERATURE REVIEW | 22 |
| 4.1. SCOPE | 22 |
| 4.2. THE RANGE OF DATA QUALITY MEASURES | 22 |
| 4.3. THE ACCURACY OF DATA RECORDING IN PRIMARY CARE | 25 |
| 4.4. PUBLISHED STUDIES ON DATA QUALITY IMPROVEMENT | 26 |
| 4.5. CONCLUSIONS FROM THE DATA QUALITY LITERATURE | 27 |
| 4.6. WIDER PERSPECTIVES ON DATA QUALITY | 28 |
| 4.7. SOCIAL FACTORS IN TECHNOLOGY CHANGE | 32 |
| 5. ANALYSIS OF THE PRIMARY RESEARCH DATA | 33 |
| 5.1. WHAT DO EXPERTS AND PRACTITIONERS IN THE FIELD THINK – AND WHY? | 33 |
| 5.2. THE PERCEIVED IMPORTANCE OF DATA QUALITY | 33 |
| 5.3. GOOD DATA QUALITY DOES NOT JUST HAPPEN | 34 |
| 5.4. PRACTICES NEED TRAINING AND SUPPORT | 35 |
| 5.5. THE DATA QUALITY FACILITATOR IS A SPECIALIST ROLE | 37 |
| 5.6. THE IMPACT OF EXISTING DATA QUALITY STANDARDS | 38 |
| 5.7. THE EFFECTIVENESS OF CURRENT MEASURES | 39 |
| 5.8. CLINICAL USERS WILL EXERCISE JUDGMENT | 41 |
| 6. CONCLUSIONS | 42 |
| 6.1. INITIAL RISKS IN THE ABSENCE OF DATA QUALITY STANDARDS | 42 |
| 6.2. ONGOING RISKS IN THE ABSENCE OF DATA QUALITY STANDARDS | 42 |
| 6.3. POTENTIAL DATA QUALITY STANDARDS: OPTION APPRAISAL | 43 |
| 6.4. HEALTHSPACE AS A MARGINAL SOURCE OF DATA VALIDATION | 46 |
| 7. RECOMMENDATIONS | 49 |
| 7.1. REMIT | 49 |
| 7.2. SUMMARY CARE RECORD | 49 |
| 7.3. DATA QUALITY VALIDATION | 50 |
| 7.4. SUGGESTIONS FOR FURTHER RESEARCH | 51 |
| 7.5. A NATIONAL STRATEGY FOR DATA QUALITY | 52 |
| REFERENCES | 53 |
| APPENDIX B: DELPHI INSTRUMENT | 57 |
Acknowledgements

We would like to extend out thanks to the following people and organisations, whose co-operation was invaluable. We would like especially to thank:

This study was funded via a research grant from the Department of Health via the Connecting for Health Evaluation Programme (CFHEP/002); the CFHEP is organised through Prof Richard Lilford’s team at the University of Birmingham. We would like to thank Connecting for Health and Sophia Layton in particular, who has been an indispensible source of information and contacts.

We thank the research participants in the Early Adopter sites who allowed us to carry out site visits and interviews and whose experiences contributed the core of the lessons learnt in this evaluation.

We would also like to thank the numerous data quality experts, clinicians and managers that participated in the focus groups and the Delphi exercise, whose insights were extremely valuable. Finally, we are very grateful to the referees whose comments were extremely useful.

The research was overseen by a multidisciplinary advisory group which included patient representation, and we are grateful for their oversight. We would also like to thank the colleagues that have made many helpful suggestions and comments on early drafts.
Preface

There are two opposing narratives on data quality in English general practice. One suggests that the quality of such data is highly variable, and that a substantial tail of poor quality in some practices threatens to jeopardise the success of the various initiatives for sharing electronic records in the National Programme for IT, notably the Summary Care Record (which formed the focus of this study). The other narrative suggests that data quality in English general practice is the envy of the world. In no other major country is there a comparable level of quality and consistency in the recording and coding of clinical data, in the capacity to produce aggregated data at community level to inform the planning and commissioning of health care, and in using routinely collected clinical data in community-based research.

Both these narratives contain an element of truth. General practice in the UK has a long tradition of considering both individual and population health; of proactive as well as reactive care; of long-term continuity of care, especially for chronic diseases; of team-based, multidisciplinary care; of a gatekeeping role for patients seeking to access secondary care; and of active participation in research. This broad and ambitious remit – the key components of a primary care led National Health Service – has driven general practitioners and their staff to build a much stronger foundation for high-quality electronic data than is the case in countries such as Germany and the USA, where ‘office based physicians’ still take a largely individualist, reactive and short-term perspective on their patients’ health. But whilst the vision, underlying values and basic infrastructure for world-class data quality are largely in place, much work remains to be done to achieve this potential.

This study addressed the question ‘what is the best way to continue to drive up the quality of data in general practice?’ We were asked to evaluate existing approaches and consider alternative ones. We encountered examples of outstandingly good practice – and of approaches that surprised and concerned us. Our remit was relative abstract (to assess the approach to data quality in general, not to pass judgement on the data quality actually found in any particular practice), and as such we were more interested in the processes by which data quality changed over time than in the static picture captured on any particular day. Broadly, we came to the conclusion that the processes currently in place are working well, that they should continue to be resourced, and that whilst pockets of poor quality data (and weak processes) are certainly evident, the future is relatively bright for addressing these problems.

Professor Trisha Greenhalgh OBE
University College London
May 2008
1. **Executive summary**

1.1. **Scope and context**

1.1.1. This report is published after a three-month evaluation by a team of academics from University College London. Our remit was to investigate the range of options for data quality (DQ) standards that can be applied to the Summary Care Record (SCR). A separate report on the general evaluation project [1] was published on 6th May 2008.

1.1.2. This evaluation used mainly but not exclusively qualitative data gathering and analysis methods, comprising around 15 interviews with NHS staff; around 30 items of grey literature, two online focus groups and a two round policy Delphi process. Data from our primary sources have largely been addressed by means of thematic content analysis.

1.1.3. Secondary data came from around 40 academic papers on data quality in health care. These were selected from over 300 candidate papers by criteria of relevance to UK primary care, relevance to data quality, relevance to electronic patient records, and date of publication. We have also examined a body of literature detailing supporting theory such as Socio-Technical Systems (STS) research, and research into the emergence of routines. These latter are not presented here, but they have informed our recommendations.

1.2. **Key findings**

1.2.1. Data quality, for the purposes of this report, refers to the extent to which a data item or items on the SCR will help with effective decision-making by users of that record. DQ is not solely a question of whether data about the patient is “right” or “wrong”, but more generally, whether that data has the potential to be misleading.

1.2.2. Effective approaches to ensuring data quality are desirable in and of themselves, if data in the SCR is to be used for effective decision-making. Good data quality measures will also permit CFH to reassure patients and clinicians of the reliability of the SCR. Finally, having known data quality measures in place will allow data users to make informed decisions as to the extent to which the data in the SCR can be relied on.

1.2.3. In the absence of a clear and limited set of use cases for the SCR, and a priority order over these use cases, it is inherently impossible to determine a single set of prescriptive data quality standards. If the potential uses of the SCR are to remain open ended, specifying prescriptive data quality standards will continue to be problematic.

1.2.4. There are, broadly, four sets of requirements for clinical data (largely tied to four groups of users). There are: clinical care uses; service planning and commissioning uses; research uses; and patient access. We chose to focus on the clinical care use scenarios, as this is the purpose of the SCR that is most widely publicised, and as we felt that this was the area in which the most immediate and direct risks existed.

1.2.5. This includes an important family of clinical use cases: those with the highest medical risk. These are the cases where data quality is most important, as the situation is likely to require rapid decision making (thus limiting the time the clinician can spend judging the reliability and accuracy of the data) and the patient is least likely to be an available source of information.

1.2.6. Existing research into clinical data quality shows that quality is highly variable between practices and between different conditions. Comparing the results of existing studies is complicated by the fact the there is no standard definition of ‘data quality’ in use. The terms ‘accuracy’ and ‘completeness’ are used in a variety of ways, often without explicit definitions. Both terms only have meaning when defined in relation to some gold standard.
A ‘gold standard’ can, in turn, only be defined with respect to the eventual use to which the data will be put.

1.2.7. Clinicians are used to dealing with data that may not be 100% complete or accurate. Some users stated that they would seek corroboration of the SCR data regardless of the level of data quality. However, the potential end users we have sampled believe that the SCR will be welcomed as an additional, but not necessarily definitive, source of information.

1.2.8. If the data held in the SCR are known, or generally perceived, to be of poor quality, then it is likely that clinicians will choose not to consult the SCR, trusting instead their local records and the patient. As well as increasing clinical risk, poor data quality is likely to result in a poor reputation and limited uptake of the SCR by end users. Thus the risks associated with poor quality data are not limited to the possibility that clinical error will increase (although that is a real possibility), but that the SCR might fail to deliver sufficient benefit to users because clinicians fail to trust or use it.

1.2.9. Most data quality standards in current use were designed for supporting individual clinicians treating individual patients. In contrast, the SCR is intended to be shared between users in a variety of different contexts. Whilst much research has been undertaken on data quality, studies to date have not addressed the specific question of the problems that arise from interpreting inaccurate or incomplete data in cross-contextual use.

1.2.10. In the absence of quantitative measure of data quality, it is still possible to make judgments about better and worse data quality. These judgments rest on the likelihood that the data on the record will lead to better (i.e. safer and more accurate) or worse decision making on the part of the user.

1.2.11. Audits such as those in the CHART system can provide quantitative measures as to the internal consistency of individual patient records, and the degree to which prevalence rates for a group of patients corresponds to the expected values given the population in general. However, neither of these is a comprehensive measure of the quality of data in the electronic record.

1.2.12. Quantitative standards are of limited use as a) they only capture data quality at one moment in time, b) they are only applicable to the minority of clinical data and c) they cannot capture the flexible and open ended requirements that the SCR must address.

1.2.13. Several observations from the focus groups, the Delphi exercise and the interviews support the conclusion the QoF is a major driver towards coding clinical data. However these same sources have also admitted what is widely known about payment-related quality metrics: data completeness efforts are often focussed on those items that have a business case, and such metrics are in no way a reflection of the generic data quality standards or data quality assurance practices of the organisation or team. Making the QoF the main driver of data quality requires only limited resources and is likely to lead to high levels of compliance with those audits.

1.2.14. Audits do provide diagnostic information about data quality in the patient record. The data they provide has indicative value as to the likely source of systematic mistakes in data production. These indicators can then be used to drive a process of organizational change that will lead to the production of good data. A combined process of audit and intervention is the most effective way to improve the clinical usability of data in the SCR.

1.2.15. Even though CHART-like audits are limited in the scope of the data items they can investigate, the data quality management problems they identify, and their solutions, have a generic benefit for the quality of the practice records. Once these problems have been identified, it is then possible to determine the source of the problems (e.g. problems with the process of data entry or coding) which can then be addressed by process improvement interventions: for example, drawing up protocols for data capture or designing and targeting staff training courses.
1.2.16. The production of good quality data relies more on the skills and professionalism of the people producing that data than it does on the systems it is recorded on. There are a number of specialist skills required in order to produce good quality data. These include technical abilities, the ability to interpret CHART results, the ability to develop good process and so on. There are also a number of professional behaviours that must be adopted in order to produce good quality data. These skills and behaviours are unlikely to arise spontaneously from within a practice. A good relationship with a skilled data quality facilitator is required in order to help practices develop these skills.

1.2.17. Whilst the initial items on the SCR are tightly defined, as the SCR grows in scope, the definition of what should be present will become more nebulous and more judgment-laden. Templates and audits can help to prevent and diagnose mistakes, but the process of ensuring data quality rests on the skills and attitudes of those individuals that are producing the data.

1.2.18. Much preparatory work was undertaken by PCTs and GP practices to improve data quality for the SCR, and this linked with wider national incentives on primary care data quality. Whilst many participating practices began with mediocre data, the process for data quality improvement was good, and rapid progress was made in many (though not all) practices. Participation in the SCR programme appeared to be a strong incentive for data quality improvement initiatives that had wider relevance to PCTs and GP practices.

1.2.19. The SCR is not the only motivating factor to improving DQ in electronic patient records. At the time that the participants in the Early Adopter phase of the SCR were being selected there were several data quality initiatives. These included: the Quality and Outcomes Framework (QOF); the Primary Care Information Service (PRIMIS) tool “Care and Health Analysis in Real Time” (CHART); the Paperlight accreditation process and the Information Management and Technology Directly Enhanced Service (IM&T DES) programme.

1.2.20. Whilst patient self management is in theory likely to encourage greater use by patients of electronic health records and therefore to enable them to correct errors, it is not yet clear what information resources in the SCR will be of value to such patients. They are already likely to know their own main medical summary data, and are more likely to value a condition specific management application which is shared with their primary and secondary care team than a summary record.

1.3. **Key recommendations**

1.3.1. **Summary Care Record**

This evaluation recommends that relevant stakeholders in the National Programme:

- continue to validate the quality of general practice data before it is uploaded to the SCR;
- publish a focussed and clearly specified set of priority use cases for the SCR;
- undertakes a formal requirements analysis of each use case, to determine the minimum utility standards that need to be met by the SCR;
- undertake a formal risk analysis of each use case to determine the minimum safety standards that the SCR data needs to meet;
- fund further investigations into the data quality requirements for summary data to be safely interpreted out of the context in which it was composed, and by diverse clinical users and patients;
• recognise that the SCR will never be 100% complete and accurate with respect to the state of the patient;

• conduct detailed evaluations of the early experience of users of the SCR, specifically inform the data quality requirements and critical safety features, as well as establish good practice in the safe and appropriate use of the SCR as a contributor to informed clinical care.

1.3.2. Data quality validation

In the absence of prescriptive DQ standards that are driven by a limited, clearly defined set of use cases, it is our position that the best possible standards of data quality will be obtained through the use of ongoing audit and intervention cycles. This is a costly approach, and requires a cohort of skilled DQ facilitators that are able to engage with the practices. Nevertheless, we would argue that the audit and intervention approach is the only way to improve the clinical usability of data in the SCR. It is imperative that interventions are carried out by DQFs that have the skills and flexibility to fully engage primary care staff in order to help them change their practices.

This evaluation therefore recommends that relevant stakeholders in the National Programme:

• continue to recognise data quality as primarily organisational and educational processes that are ongoing and require regular stimulus;

• extend the availability of well-trained Data Quality Facilitators to assist practices to appraise their data quality; identify gaps and provide training in data quality improvement and monitoring; share good practice approaches;

• continue to support organisations and tools used to audit data completeness and accuracy in primary care, such as PRIMIS+ and CHART, and the Paperlight protocol;

• recognise the limited "indicator" role that established metrics of data quality play in implying but not establishing the general level of quality of practice health records, serving more as an indication of processes and training needs than an assessment of the fitness of the records for CRS upload;

• recognise the limitations on the role of HealthSpace as a route to data quality validation;

• fund new research to identify suitable evidence based quality processes and data quality metrics to establish the fitness of a practice and its records to support the defined purposes of the SCR.

1.3.3. A national strategy for data quality

Given the involvement of multiple stakeholders in the SCR programme, and the complexity of determining the responsibilities of these, we strongly recommend that a national strategy for data quality should be established.
2. Introduction

2.1. Remit

2.1.1. CfH commissioned this evaluation in order to determine a) the effectiveness of data quality measures in place at the moment and b) the data quality measures that are desirable in order to ensure the widespread use of the SCR. Data quality is undoubtedly a prerequisite for producing a meaningful SCR, but there is uncertainty as to whether current standards are fit for purpose. Most data quality standards in current use were designed for supporting individual clinicians treating individual patients. In contrast, the SCR is intended to be shared between users in a variety of different contexts. Whilst much research has been undertaken on data quality, studies to date have not addressed the specific question of the problems that arise from interpreting inaccurate or incomplete data in cross-contextual use.

2.1.2. Ensuring data quality is a challenge in any data sharing exercise, irrespective of the data being collected. Data quality for the SCR is especially problematic for several reasons: firstly, the risks from poor quality data, especially in the clinical use setting, are potentially very high. Secondly, the SCR will be populated with data that is drawn from the primary care setting. The field of primary care does not always lend itself to the recording of firm diagnostic codes: much of primary care illness is less precisely formulated. Whilst GPs accept that secondary users (i.e. researchers and planners) require coded data, primary care problems don’t often fit into this biomedical model [2]. Finally, if the SCR has, or is perceived to have, low data quality, user uptake will be low. With low user uptake, few of the hoped-for benefits of the SCR will be realised. It is, therefore, imperative that there be measures to identify and rectify problems with data quality.

2.1.3. Effective approaches to ensuring data quality are desirable in and of themselves, if data in the SCR is to be used for effective decision-making. Good data quality measures will also permit CfH to reassure patients and clinicians of the reliability of the SCR. Finally, having known data quality measures in place will allow data users to make informed decisions as to the extent to which the data in the SCR can be relied on.

2.1.4. Our remit is to provide useful insights to CfH in order to assist with the drawing up of data quality standards. The pragmatic questions asked by the funder are:

i. “If no data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records occurred, what would be the risks in terms of the completeness and accuracy of the derived summaries?”

ii. “If no data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records occurred, would this have implications for the subsequent upload of problems and diagnoses.”

iii. “If it was considered desirable that minimum data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records, what might those data quality standards be in order to maximise patient benefit and minimise patient risk?”

2.1.5. In order to answer these questions we have examined a range of theory and evidence. From secondary and primary data we have determined:

i. What is already known about the benefits of good data quality and the hazards of poor data quality in relation to networked electronic patient record systems?

ii. What approaches to data quality standards have been tried in other projects within and outside the UK, and what were the critical success factors and key drawbacks of these (taking note of the type of data shared in each case)?

iii. What experts and practitioners in the field think – and why?
iv. What impact the existing data quality standards have had on the Early Adopter SCR implementation, and what lessons can we learn from this?

2.1.6. In the following sections we will present our findings and our recommendations. Firstly we will discuss the methods adopted during this research. We discuss the problem of circumscribing the contexts of use for the SCR. We will present a number of illustrative use cases in order to highlight the range of potential data uses and users. We will discuss a body of literature which will address a number of questions about existing data quality measures. We will present the key findings from our primary research data. Our conclusions, drawn from this material, will be elaborated on and we will conclude the report with an overview of the advantages and disadvantages of these options and a set of recommendations.

2.2. **Background to data quality concerns in relation to the SCR**

2.2.1. There are, broadly, four sets of requirements for clinical data (largely tied to four mainstream groups of users). There are: service planning and commissioning uses; research uses; clinical care uses; and patient access. We chose to focus on the clinical care use scenarios, as we felt that this was the area in which the most immediate and direct risks existed. This is because:

   i. In the research setting, and in the planning and commissioning settings, risks from poor data quality are relatively low. When dealing with population data it is always possible to place confidence intervals around the aggregated data obtained. The higher the quality of data, the tighter those confidence intervals would be. Nevertheless, even fairly approximate data would have some indicative value and could lead to the development of useful models. Judicious use of such data can still support good planning and research outcomes.

   ii. In the patient access setting, the risk centres around the credibility gap that would arise, should the patient see mistakes (real or perceived) in their SCR. The more categories of information on the SCR, the more scope there is for error. There is a risk of this credibility gap arising, not only if data do not accurately reflect the patient’s real record but also if the SCR and HealthSpace do not meet the expectations of the patients. Whilst this may affect the patient’s level of trust in the competence of their clinicians, their health is unlikely to be directly affected.

   iii. In contrast, in the clinical care setting, confidence intervals cannot be applied to individual data items in a patient’s record. An individual patient record must contain a data item, or a ‘not known’, for each of the headings in the SCR. The risks of acting on an inaccurate record are potentially very grave. For this reason, we conclude that the highest priority in ensuring data quality should be in determining its fitness for the purposes of clinical care.

2.2.2. The drive to improve data quality in primary care, and the potential benefits associated with this challenge, occurred on a much wider canvas than the specific SCR programme. Before the SCR programme began, many PCTs had already appointed their own ‘generic’ data quality facilitators (DQFs) to support IT development in GP practices, and CfH also provided specific DQFs for the SCR programme. It is also worth noting that the existence of the SCR, and participation in the SCR programme, appeared to provide a strong incentive to PCTs to work on improving their data quality.

2.2.3. The main data quality standards in use in GP practices at the time that selection for inclusion in the Early Adopter SCR programme was occurring were:

   i. The Quality and Outcomes Framework (QoF), a national system of financial incentives and rewards for various clinical and administrative targets in general
practice. To win QoF points, practices are required to collect data on both the process and outcome of care across different disease groups and dimensions of management. There is some evidence that the QoF has begun to have a positive impact on patient outcomes [3].

ii. Primary Care Information Service (PRIMIS+, www.primis.nhs.uk/), a national programme of training and quality improvement for GP practices in relation to ICT. PRIMIS is an academic group based in Nottingham, which has a recurrent contract with the Department of Health. The group has produced a series of semi-automated audits known as ‘CHART’ (Care and Health Analysis in Real Time), which have proved popular and credible amongst GP practices and which are widely used as tools for data quality improvement initiatives (and also by PCTs to monitor practices’ progress in this sphere). An interactive example of a CHART audit is accessible from the PRIMIS website (http://www.primis.nhs.uk).

iii. Paperlight, an accreditation process that ensures GP surgeries are able to work exclusively with computerised patient records. In order to qualify for Paperlight accreditation, practices must have: complete patient records on the computerised clinical system; policies to ensure those records are maintained; regular audits to validate those records; a software system that is accredited under GPSoC (GP Systems of Choice, a limited selection of approved software systems for use in GP practices); data security measures and business continuity provisions.

iv. The IM&T DES (Information Management and Technology Directly Enhanced Service) was a dedicated funding stream negotiated nationally by the National Clinical Leads that provided incentives to GP practices to put in the major effort needed to improve data quality. The IM&T DES ran from 1st April 07 and is to be supported until 31st March 2009.

2.2.4. Data on the SCR are, at present, drawn from a single data source: the GP held Local Detailed Record (LDR). If data on the LDR are poor then the SCR will be inaccurate, incomplete and therefore untrustworthy. Data quality is undoubtedly a prerequisite for producing a meaningful SCR, but there is uncertainty as to whether current standards are fit for purpose. Most data quality standards in current use were designed for supporting individual clinicians treating individual patients, and with the assumption that the full computerised (local) record will be available. In contrast, the SCR is intended to be shared between users in a variety of different contexts, and as the only available extract of the non-local records for a patient. Whilst much research has been undertaken on data quality, studies to date have not addressed the specific question of the problems that arise from interpreting inaccurate or incomplete data in cross-contextual use.

2.2.5. Few clinical records (paper or electronic) are 100% complete or accurate, and it is an established part of clinical practice to work on the assumption that data may be incomplete or inaccurate, regardless of source. For example, just because a patient has no recorded allergies doesn’t mean they have no allergies and clinicians will often check with the patient regardless of the content of the record. If the data held in the SCR are known, or generally perceived, to be of poor quality, then it is likely that clinicians will choose not to consult the SCR, trusting instead their local records and the patient. Thus the risks associated with poor quality data are not limited to the possibility that clinical error will increase (although that is a real possibility), but that the SCR might fail to be of benefit to patients or to users because clinicians fail to trust or use it.

2.3. Defining data quality standards

2.3.1. Previous research on data quality standards tends to discuss the completeness and accuracy of data. However, it is impossible to define an absolute meaning for these terms:
data can only be defined as ‘complete’ or ‘correct’ with respect to some particular use of the data. Studies have been unsystematic in their application of these terms, making comparison difficult [4].

2.3.2. The terms complete and correct can only be defined with respect to some gold standard or particular use of the data. It follows that, in the absence of narrowly defined use cases for the SCR, it is not possible to determine a single set of prescriptive data quality standards. If the potential uses of the SCR are to remain open ended, specifying prescriptive DQ standards will be problematic. At the time of writing, there remain two major uncertainties relevant to data quality:

i. What will the main (and the potential supplementary) uses of the SCR?

ii. What information will the SCR eventually contain?

In the absence of a clear and limited set of use cases for the SCR, and a priority order over these use cases, it is inherently impossible to determine a single set of prescriptive data quality standards.

2.3.3. However, even in the absence of prescriptive data quality standards, it is possible to make reasonable statements about ‘better’ or ‘worse’ data quality. If good data are defined as those which are ‘fit for purpose’, for example, then a working definition of poor quality data might be ‘data that could mislead a clinician in one or more contexts of use’, and a working definition of good quality data might be ‘data that lead to better decision making in one or more contexts of use’. Note that these definitions, whilst pragmatic and defensible even in the absence of specific use cases, do not lead directly to quantifiable measures of data quality.

2.3.4. Audit measures, such as the ones in the CHART system, do result in quantified outcomes. Queries of the LDR can determine where data are possibly absent or materially incorrect, for example where a record of a prescription (e.g. digoxin) is present but the appropriate diagnostic code (e.g. atrial fibrillation) is not. These audits are not, in and of themselves, sufficient to ensure data quality. Firstly, it is only possible to generate audit queries of the ‘drug-diagnosis correlation’ type for a limited subset of data items. For most clinical data, no meaningful rule based queries can be constructed. Secondly, because the current plan is for the SCR to be used in a variety of settings, it is not possible to determine a priori which data need to be audited. Finally, audits such as the CHART queries only reveal the state of one particular set of data at one particular moment in time.

2.3.5. These audits can, however, point to underlying problems with the data capture process, such as problems with the process of data entry or coding. Once these problems have been identified, it is then possible to determine the source of the problems that can then be addressed by process improvement interventions: for example, drawing up protocols for data capture or designing and targeting staff training courses. Even though CHART-like audits are limited in the scope of the data items they can investigate, the data quality management problems they identify, and their solutions, can have a generic benefit for the quality of the practice records.

2.3.6. GP practices do not necessarily have the skills to carry out process improvement interventions. Even where practices have staff that are skilled in both IT and data quality, they may need support in carrying out CHART and other audits, analysing them, and identifying which of the practice’s processes and procedures to target for improvement. To this end, the role of the data quality facilitator (DQF) has been developed.

2.3.7. Our research has shown that, where DQFs are remote, inflexible or lack credibility, their interventions will not suffice to improve data quality. Where they work well, DQFs are flexible enough to adapt their expertise to the needs of different practices. Our findings show that in these cases DQFs are very effective at improving data quality.
2.3.8. In the absence of prescriptive DQ standards that are driven by a limited, clearly defined set of use cases, it is our position that the best possible standards of data quality will be obtained through the use of ongoing audit and intervention cycles lead by a DQF. This is a costly approach, and requires a cohort of skilled DQFs that are able to engage with the practices. Nevertheless, we would argue that the audit and intervention approach is the only way to improve the clinical usability of data in the SCR.

2.4. **Content of the report**

2.4.1. The subsequent sections of this report document: the research methods employed; the range of the contexts of use for the SCR; an overview of the literature addressed; an analysis of the primary research data we have collected; the conclusions and key recommendations. Supplementary material is contained in the appendices.

2.4.2. For the sake of consistency we will use the following terminology throughout this report. A "gold standard" is a source of data about the patient that provides a basis for comparison for the electronic patient record. Gold standards include prescription records, paper records or an evaluation of the patient for example. Gold standards are not necessarily infallible but they provide a useful comparator for the electronic record. The completeness and correctness of records can only be determined with respect to a gold standard. Throughout we will use the Hogan and Wagner definitions of these measures [5]. Completeness is the proportion of a group of patients’ conditions, medications and observations that exist in a gold standard record that are also present on the record. Correctness is the proportion of conditions, medications and observations that are in a group of records that are correct with respect to the gold standard. Accuracy is a hypothetical measure of how well the record reflects the reality of the patient condition. This can not be objectively measured but is frequently used within CfH and the early adopters, and in the wider literature, as an informal quantity. Correctness is a proxy for accuracy, and can be measured.

2.5. **Methods**

2.5.1. This evaluation used mainly but not exclusively qualitative data gathering and analysis methods, comprising around 15 interviews with NHS staff; around 30 items of grey literature, two online focus groups and a two round policy Delphi process. Data from our primary sources have largely been addressed by means of thematic content analysis.

2.5.2. Secondary data came from around 40 academic papers on data quality in health care. These were selected from over 300 candidate papers by criteria of relevance to UK primary care, relevance to data quality, relevance to electronic patient records, and date of publication. We have focused on:

- **Systematic reviews of data quality probes.**
- **Quantitative and qualitative studies of the impact of the introduction of clinical electronic record and ordering systems.**
- **Qualitative reviews of data quality practice.**

We have also considered the body of literature detailing supporting theory such as Socio-Technical Systems (STS) research, and research into the emergence of routines. These have informed our key recommendations but we do not present the detailed literature in this report.

2.5.3. In order to better understand the background to the SCR in particular, and data quality in general, we have carried out desk research on a number of items of Grey literature including:

- **Summarising protocols**
• Training and information literature
• Data Quality Facilitator Job Descriptions
• GP practice internal audit templates
• A number of internal reports
• External reports including the BCS evaluation [6]

2.5.4. We have gathered a substantial amount of primary data. We have sought input from two main groups: data quality experts, and clinicians and supporting staff. The clinicians and supporting staff can further be broken down into two groups: data producers (GPs, practice managers, practice IT leads etc.) and data users (out of hours and acute clinical staff, community nurses etc.) For each of these groups we have selected participants from the UK and overseas.

2.5.5. We ran two online focus groups, one consisting of GPs and Practice managers and one consisting of Data Quality experts. These focus groups consisted of structured questions that were prompts for discussion. Participants were free to contribute at any time over several weeks.

2.5.6. We also completed a two-round online policy Delphi process. A Delphi process consists of multiple “rounds”. In each round the participants are presented with a set of statements or options and asked to rate them along a Likert scale. Participants are also asked to explain the reasons for their choice. In rounds two and beyond, participants are presented with their own responses, plus the aggregated data from the other participants’ responses. The aim of many Delphi exercises is to reach consensus, as the participants integrate the other respondents’ points of view into their own responses [7]. A policy Delphi is structured so as to avoid consensus however, as the aim is to identify the range of advocacy positions taken by the participants [8]. In our Delphi exercise we received input from:

▪ A&E Consultants
▪ GPs
▪ Out of Hours and Walk in Centre Nurses
▪ Nurses from NHS Direct
▪ Data Quality experts within CfH and PRIMIS
▪ International experts on clinical data quality

We have Carried out interviews with ~15 personnel from early adopter practices, including:

• IT Leads
• Practice Managers
• GPs
• Nurses (Practice and Walk in Centre)
• Summarising personnel

2.5.7. We also carried out a series of interviews. Our interviews took a variety of formats, including site visits, telephone interviews and interviews at UCL. Interviews began in an unstructured fashion in order to determine the topics and issues that were of greatest priority to the interviewee. Before closing, specific questions were asked in order to elicit particular pieces of information. Notes were made contemporaneously and typed up immediately. Tape recorders were not uses, so as to encourage the free discussion of sometimes sensitive issues.
2.5.8. We analysed the Delphi, focus group and interview data by means of thematic content analysis. Themes were selected from the review of the literature, the discussions within the focus group and from preliminary interviews. These themes were then used as the basis for the design of the first round of the Delphi process, and for the planned questions in later interviews. These same themes were used to analyse the primary data. However, further themes arose during the final analysis. Box 1 shows the themes that were addressed.

2.5.9. The data we obtained were uncontroversial in that there were no marked disagreements between participants. Different groups of respondents had slightly different foci to their responses but no contradictions in positions were unearthed. Furthermore, the data we uncovered was consistent with the literature that we have studied.
3. Use case scenarios for the SCR

3.1. The range of use cases

3.1.1. We have already alluded to the difficulties that arise due to the fact that the uses of the SCR remain undefined. In their 2006 report, the British Computer Society stated that:

"Arguably the major weakness of NHS CFH is that it currently lacks a business context... The output-based specification (OBS) used in the NPfIT tendering process is not – and never was – a substitute for business requirements... Business objectives should drive information objectives which in turn should drive IT solutions." [6]

Use case scenarios document the range of ways in which a specific user may interact with a system. A set of Use Case Scenarios should have been developed as part of the initial business plan. These would serve to focus the scope of the SCR, and to clarify what is and is not required of the system that is eventually put in place.

3.1.2. At present there is an open-ended range of users and requirements for the SCR, some of which are mutually contradictory. It is therefore impossible to devise a set of prescriptive data quality standards as levels of detail and sets of data items differ according to these requirements. For example, users in emergency medicine settings may need to know whether the patient has a tendency towards high blood pressure. If this is the highest priority use case then the SCR should contain a single entry detailing whether the patient is hypertensive or pre-hypertensive. Community nurses on the other hand may need detailed trend data about the progress of the patient’s blood pressure measures over recent weeks and months in order to determine whether the patient’s self management is or is not effective. Without knowing how the SCR is to be used, and more importantly, what its primary uses are, it is impossible to determine which of these data definitions is the most appropriate.

3.1.3. Tables 1 to 4 show a number of example use cases for the different classes of user. This table is by no means exhaustive, but it demonstrates the range and variety of requirements that end users may have.

3.2. Use cases in the clinical setting

3.2.1. Table 1 gives examples of several Use Case Scenarios in the clinical setting. These are only illustrative examples. The range of settings includes those in the unscheduled acute, unscheduled chronic, unscheduled minor and scheduled but remote settings. Each of these sets of users potentially has several Use Case Scenarios. These scenarios vary from low to high risk. Risk is higher wherever:

1. secondary sources (including the patient or carer) against which data can be verified are unavailable
2. the decision to be made is urgent
3. the data being accessed is of central to the clinical judgement to be made or
4. the data will be used in an automated system, rather than being subject to professional judgement.

Therefore the various settings may be more or less likely to experience these factors that raise risk, but risk is not tied to setting per se.

3.2.2. Data quality must be addressed in order to minimize this risk. However, the requirements for data quality depend not only on the level of risk tied to the use of that data. They also
depend on the use to which the data will be put, as different uses lead to different sets of required data, as in the example of appropriate blood pressure readings for community nurses versus emergency care clinicians. Thus the range of potential uses leads to a variety of conflicting requirements and standards for data quality in the clinical setting.

### Table 1: Clinical Use Case Scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Use of SCR</th>
<th>Most important DQ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>To provide unplanned urgent care to a patient with no (recent) local records, and who is unable to provide this history.</td>
<td>View conditions and current medications, to help establish possible causes and contributors to current problem</td>
<td>Completeness and accuracy of condition list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness and accuracy of medication items list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent test results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent/forthcoming clinic appointments</td>
</tr>
<tr>
<td>To provide unplanned urgent care to a patient with no recent records: deterioration or complication of a long-term condition</td>
<td>View current medications, conditions list less vital but ideally any recent consultations with other care providers</td>
<td>Currency, completeness and accuracy of medication dosages</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent test results</td>
</tr>
<tr>
<td>To verify or elaborate the details of a patient’s medical history, given orally, in the absence of health records e.g. NHS Direct tele-consultation, home visit by deputy</td>
<td>View conditions, medications, family history, test results, appointments, ideally all recent consultations</td>
<td>Accuracy of condition list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accuracy of medication items list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent test results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent/forthcoming clinic appointments</td>
</tr>
</tbody>
</table>

3.2.3. Table 2 gives examples of a number of Use Case Scenarios for the patient user of the SCR. The patient may wish to use the SCR in order to check their understanding of, or refresh their memory about, details of their clinical care. They may wish to use the SCR to check whether the data on their record accords with their understanding of their clinical care, or they may wish to use the data on the record to support their own self-management of their condition.
### Scenario Use of SCR Most important DQ criteria

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Use of SCR</th>
<th>Most important DQ criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>To check on the accuracy and completeness of the summary</td>
<td>View all content. Ensure that no mistakes have been made, or that embarrassing (correct or incorrect) information is not present</td>
<td>No erroneous data (esp. conditions, family history)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness of conditions list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness of medications items list</td>
</tr>
<tr>
<td>To confirm the outcome of a recent admission or consultation; to verify understanding of what was told orally by the care providers.</td>
<td>View conditions list, medications</td>
<td>Currency of conditions list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Currency of medications and doses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent test results</td>
</tr>
<tr>
<td>To share records with a carer or relative: to support their self-management of their care</td>
<td>View all content, concern that it contains no mistakes, and is up to date</td>
<td>Accuracy of data (esp. conditions, family history)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Currency of conditions list</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Currency of medications and doses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any recent test results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any forthcoming clinic appointments</td>
</tr>
</tbody>
</table>

### Table 2: Patient Access Use Case Scenarios

3.2.4. The information seeking and self-management scenarios require good quality data. If the patient is to act on this data then it must be acceptably complete and correct. Where the patient is validating their record, data quality is still important, though the risk becomes one of a loss of credibility in the patient’s own record and, potentially, the clinicians that produced it.

3.2.5. However, the use of HealthSpace is complicated by various factors such as patient disengagement and the current level of involvement required in order to obtain a HealthSpace account. (See [1] for more details).

3.2.6. Table 3 outlines several example use cases for the research setting. The risks arising from the research use of the SCR are relatively low. The data will be used for indicative purposes only in these scenarios, and should be verified with the research subject and their clinician before any intervention takes place.

### Table 3: Research Use Case Scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Use of SCR</th>
<th>Most important DQ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>To screen patients for inclusion/exclusion diagnoses for a clinical trial</td>
<td>View population queries and anonymised extracts: conditions</td>
<td>Completeness and accuracy of conditions list must be known</td>
</tr>
<tr>
<td>To assess population incidence of a diagnostic profile, prior to designing a study</td>
<td>View population queries and anonymised extracts: conditions</td>
<td>Completeness and accuracy of conditions list must be known</td>
</tr>
</tbody>
</table>

3.2.7. Commissioning and planning use scenarios are also relatively low risk. Table 4 gives examples of commissioning uses. As commissioning and planning are prospective activities, and as the data used for commissioning and planning purposes are always historic, there is always the potential for data used in these settings to be outdated and
therefore misleading. Planners are well used to dealing with trend information that may be more or less reliable. It is commonplace to consider margins of error in these cases.

Table 4: Commissioning and Planning Use Case Scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Use of SCR</th>
<th>Most important DQ Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>To compare local and national incidence and/or prevalence of particular conditions</td>
<td>View geographic sub-population queries: conditions</td>
<td>Completeness and accuracy of conditions list must be known</td>
</tr>
<tr>
<td>To profile and monitor drug utilisation (for costing and planning purposes)</td>
<td>View geographic sub-population queries: medication items</td>
<td>Completeness and accuracy of medication items list must be known</td>
</tr>
</tbody>
</table>

3.2.8. In the planning and commissioning setting, the data will be dealt with at population level. Consequently it will be possible to determine the margins of error that apply. The most pressing requirement as far as data quality is concerned is to determine the level of reliability of the data that is available. This will allow the users to determine the appropriate weight of evidence to accord to the records.

3.3. Delphi data concerning use case priorities

3.3.1. We carried out a two round policy Delphi, as described in the Methods section. The Delphi instrument is reproduced in Appendix B. The main results are analysed by theme in Section 5. In order to circumscribe the range of clinical uses we have considered, we here present data concerning the use cases that respondents felt that the SCR might fulfil. Respondents were asked to indicate the priority that they would accord to the clinicians in each vignette being able to access the SCR. Respondents were asked to rank the importance of the SCR in each of these vignettes in terms of relative priority in order to encourage them to consider the relative benefits and risks in each case.

3.3.2. Respondents were asked to consider eight vignettes at two levels of risk, as measured by the level of completeness and accuracy of the SCR with respect to the medical history of the actual patient. In the low risk (high quality) setting, there was a 90% probability that the patient’s record would be complete and a 98% probability that the data in that record would be correct. In the higher risk (medium quality) setting there was a 70% probability that the patient’s record would be complete and a 90% chance that it would be correct. Results are presented in Table 3 and Table 4.

3.3.3. The first item of note is that the priorities for use did not change a great deal at different levels of data quality. Rather respondents, particularly those with a clinical background, suggested that the weight accorded to the evidence in the SCR would diminish, but the use cases would not change.

3.3.4. If there was clear agreement as to the most desirable use cases for the SCR we would expect to see large “clusters” of responses around one priority for each of the vignettes. Perhaps surprisingly, the priorities over the all but one of vignettes were not clear-cut. This indicates that there is a wide range of opinion as to the most important roles that the SCR will play. The junior doctor vignette was considered to be one of the top three priorities to use the SCR by 14 respondents at both levels of data quality (and in the bottom three

---

1. In the Delphi instrument we used the terms complete and accurate rather than complete and correct, and did not introduce the notion of a “gold standard”. Despite the imprecision, we chose to use these terms because “complete” and “accurate” are in common use in CfH and among the earlier adopters and it was felt that the respondents would be most comfortable considering the records in this manner.
priorities by only two). The district nurse and NHS direct vignettes were around twice as likely to be one of the bottom three priority than to be rated one of the top three. There was still a great deal of variance in the priorities accorded to these vignettes by the respondents. The other three vignettes (walk in centre nurse, triage nurse and GP) were all about as likely to be rated as one of the top three as one of the bottom three priorities.

3.3.5. In general, it seemed that the vignettes themselves did not prompt the respondents to generate clear-cut priorities over these use case scenarios. Where respondents gave reasons for their choices, the nature of the users themselves, and the clinical elements of the vignette, did not account for much of the determination of the priorities.

Table 5: Priorities where the SCR is 90% complete and 98% accurate

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Priority 1</th>
<th>Priority 2</th>
<th>Priority 3</th>
<th>Priority 4</th>
<th>Priority 5</th>
<th>Priority 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior Doctor</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Walk in Centre Nurse</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Triage Nurse</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>GP</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>District Nurse</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>NHS Direct</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 6: Priorities where the SCR is 70% complete and 90% accurate

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Priority 1</th>
<th>Priority 2</th>
<th>Priority 3</th>
<th>Priority 4</th>
<th>Priority 5</th>
<th>Priority 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior Doctor</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Walk in Centre Nurse</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Triage Nurse</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>GP</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>District Nurse</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>NHS Direct</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

3.3.6. Respondents were asked to elaborate in free text the reasons why they had selected the particular priorities for each of these vignettes. Respondents were much more likely to give reasons for low priorities than for high ones. The responses fell into five main themes, which are summarized in Table 7.
Table 7: Reasons given for down-rating the priority of using the SCR

<table>
<thead>
<tr>
<th>Reason</th>
<th>Number of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>The immediate clinical signs are more important</td>
<td>6</td>
</tr>
<tr>
<td>Data are more immediately available from patient/carer</td>
<td>6</td>
</tr>
<tr>
<td>Data may not be current</td>
<td>3</td>
</tr>
<tr>
<td>The data required is not on the SCR</td>
<td>2</td>
</tr>
<tr>
<td>More urgent tasks require attention</td>
<td>2</td>
</tr>
</tbody>
</table>

3.3.7. Priorities over the vignettes were not chosen according to the role of the user (not mentioned by any of the respondents) nor by the clinical aspects of the vignette in isolation. Instead, the factors that were mentioned most were related to the procedural issues involved in using the SCR. This accounts for much of the variance between the priorities selected: respondents had very different views of the procedure that must be undertaken by the clinician in each of these vignettes. The predominant reasons given for not using the SCR were that either: the observation of the patient was sufficient, or more important, in determining the right course of action or; the data could be obtained from the patient or carer in person. GPs in the focus group backed up this impression. The following are illustrative quotes:

"The SCR would not really be useful for straightforward, uncomplicated cases [but where the patient is] elderly, lives alone and has lots of medical conditions and therefore usually have a polypharmacy" the SCR would be useful.

"[it's] mainly useful in a setting where the clinician has no information about the patient"

3.3.8. One respondent, a Canadian Emergency Medicine expert, noted that, even where data on the SCR is potentially relevant, time spent checking the SCR is time spent away from other tasks: "seeking the [medical] profile should NOT distract from the main overall priority – the ABCs of resuscitation/emergency management". A former GP was of a similar opinion: "Acute care doesn't need 'old' information. The acute care situation can be assessed without much prior knowledge."

3.3.9. The reasons given for up-rating the priority of the junior doctor scenario were that the scenario required an urgent response (8 respondents) or to avoid an adverse reaction to treatment (4 respondents). Six respondents stated that information should be validated from other sources, regardless of the accuracy or reliability of the SCR. Common comments were" "delays are potentially life threatening", "[There is] potentially a very high need for immediate action", "Dangerous wrong decision to be avoided". This identifies an important family of clinical use cases: those with the highest medical risk. However, these are the cases where data quality is most important, as the situation is likely to require rapid decision making (thus limiting the time the clinician can spend judging the reliability and accuracy of the data) and the patient is least likely to be an available source of information.
4. Literature review

4.1. Scope

4.1.1. A range of studies concerning data quality have been carried out and are described in the literature. A literature review was performed in order to identify approaches to the assessment and improvement of data quality that have some evidence of validity and success.

4.1.2. Secondary data came from around 40 academic papers on data quality in health care. These were selected from over 300 candidate papers by criteria of relevance to UK primary care, relevance to data quality, relevance to electronic patient records, and date of publication. We have analysed this data in order to answer three main questions:

   a. To what extent do data quality measures of different sorts address the quality of data in primary care?
   b. What is the accuracy of data recording in primary care?
   c. What data quality interventions have so far been attempted, and with what success?

4.2. The range of data quality measures

4.2.1. Existing approaches to analysing data quality in clinical records can be classified under three headings: comparison with gold standards; probes of internal consistency and; comparison with external prevalence data. However, there is some overlap between these approaches.

4.2.2. Gold standards compare the patient record to some external source of data, such as prescription registries, billing data or paper records. This may be done by means of automated queries, where the gold standard data is stored in electronic form.

4.2.3. Probes are queries that are run on patient records in order to verify whether these records are internally consistent. Whilst the same types of data may be included in these queries as in the gold standard approach (for example, data about tests or prescriptions), crucially this data comes from the patient’s own record.

4.2.4. Comparisons with external prevalence data also consist of queries. In this case the queries group information from the entire register, rather than individual patients’ records. The prevalence of certain conditions is compared against external data about their prevalence in other practices. This last category is a comparison between the practice sub-population and an appropriate wider population.

4.2.5. The gold standard approach has been widely applied in studies of data quality in UK clinical practice [4]. However, the definitions of accurate data vary between studies. Hogan and Wagner proposed standard criteria [5] such that the data quality individual records, and registers of many records, will be measured in terms of completeness and correctness with respect to a gold standard. Completeness is defined as the proportion of observations about the world (that are present in the gold standard data) that are present in the patient record(s). Correctness is the proportion of observation in
the patient record(s) that represent the state of the world, as recorded in the gold standard data.

4.2.6. A systematic review ([4]) analysed studies that apply the measures of completeness and correctness to either individual records or registers of many records. The inclusion criteria for this review were that: studies to be reviewed must be i) in the primary care setting, ii) assessments of data quality in computerised records, iii) based in the UK and iv) assessing the completeness and correctness of patient records or registers. The authors excluded studies i) that used prevalence rates as a gold standard, ii) in which the patient was reassessed by another clinician, iii) that used fictional patients and iv) studies in which patients self-reported on their own medical history. In all, 24 studies were reviewed.

4.2.7. The studies reviewed in [4] demonstrate that:

a. Data completeness was highly variable between practices (completeness rates between 53% and 89% for diabetics in different practices in a single study).

b. Data correctness and completeness varied systematically between conditions, with those having clear diagnostic features (diabetes, cancer) being more likely to be correctly and completely recorded than those with more subtle clinical indications (asthma, anorexia nervosa, hypertension).

c. In the absence of interventions, there has been no clear improvement over time: studies in the early 2000s show no higher rates of completeness and correctness than studies from the 1980s.

4.2.8. The critical success factors identified by Jordon, Porcheret and Croft in determining whether practices achieved high levels of completeness and correctness were twofold: the clarity of the diagnostic criteria of the conditions being recorded and whether the practice has access to training.

4.2.9. Many of the studies reviewed in [4] acknowledged a key shortcoming: practices that took part in these studies were largely those that had pre-existing interest of training in recording information electronically. Furthermore, several multi-practice studies discarded underperforming practices from their study. As a result, completeness and correctness rates in these studies are likely to be higher than a representative average.

4.2.10. Similar gold standard studies have been carried out beyond the UK [9]. is a Belgian study that compared patient records with billing data. Whilst this study was intended to measure the quality of the billing records rather than the patient records, it still provides a comparison between patient records and an external “gold standard”. The study looked specifically at recording of three conditions: hypertension, heart failure and pneumonia. They found conditions recorded without treatment (correctness errors) and treatments recorded without conditions. These latter mistakes are potentially completeness errors but they may also be due to over-billing. The Belgian study confirmed the findings of the UK systematic review regarding the importance of clear diagnostic criteria, in that error rates were variable and that heart failure and hypertension tended to be less completely recorded than pneumonia.

4.2.11. Another measure widely used in the literature is the Probe. Probes are usually run on a set of electronic records in order to identify data that are internally inconsistent. An example of the use of these internal probes to validate the data quality in three practices is given in [10]. A set of business rules was used to drive the creation of a set of queries. For example, the business rule “unopposed oestrogens should not be
prescribed to a patient who has not had a hysterectomy” is translated into the query “identify all cases prescribed unopposed oestrogens not recorded as having a hysterectomy”. These business rules were drawn up by the “quality circle task group” in response to the trust’s primary care group quality initiatives.

4.2.12. Figure 1 shows that rates of errors ranged from low (no patients under 12 were shown to be on tetracyclines) to high (prescriptions of co-danthramer to patients with no diagnosis of a malignant tumour). No interventions were suggested in this study, but follow up queries showed that not all probe errors were data quality errors. For example, of the errors made in response to probe 1, only some of the errors were due to a failure to record a known hysterectomy (45.5% of the errors in Practice A, 79% in Practice B and 16.7% in Practice C). The authors do not discuss the reasons why practices vary in the completeness of their recordings of hysterectomies.

4.2.13. As well as demonstrating the variability of data quality across practices, this result is illustrative his result highlights a potential drawback of using probes to measure data quality. Not all errors unearthed by the probes are due to poor completeness or correctness. Data probes are not good at identifying unusual cases, such as the existence of separate prescriptions for oestrogen and progesterone. Secondly, as Figure 1 shows, data quality probes can only be constructed where a business rule exists that can be translated into a query.

4.2.14. The final method of data quality measurement identified in the literature is the comparison of prevalence rates in a practice register with external prevalence rates. PRIMIS+ publish prevalence rates for the UK based on abstraction from 20 million patient records [11]. The Royal College of General Practitioners also publishes a register of morbidity statistics from general practice (MSGP). These registers provide “rates per thousand” for a variety of conditions. For example, between 2003 and 2007,
the rate of Angina (or other chronic IHD) was 29.65/1,000 patients (24.48/1,000 for females and 34.89/1,000 for males). A practice can count the number of records with the relevant diagnosis codes and compare this against the expected outcome given the national prevalence rate.

4.2.15. Whilst prevalence rates can give a useful approximation of the completeness of a practice’s register, there is a drawback with using prevalence rates to drive data quality improvement [12]. describes interventions that used prevalence data as part of the drive to improve data quality. However, little attention is paid to how closely practice prevalence rates should approximate the national mean. In several cases the practice prevalence rates far exceeded the national prevalence rate (for example, a rate per 1,000 of 109 for diabetes mellitus codes, versus the MSGP4 rate per 1,000 of 41.2 for this condition). In the absence of measures of variance, prevalence rates do not give an indication of the acceptable range of recording rates. Furthermore, practices vary widely in the demographics that they serve: a university campus practice may be expected to have a far lower rate of age-related conditions than average for example.

4.3. The accuracy of data recording in primary care

4.3.1. Audits of data quality have consistently shown that there is wide variance between the data quality recorded in different practices ([12],[4]) and between different morbidities ([9], [4],[10] see Table 8). Practices tend to be better at recording some types of data than others [4] and very few practices were either poor or excellent across the board. For common morbidities and drug regimens (e.g. asthma patients, HRT recipients) the rate of errors in completeness was tended to be highest for recording of observations (>50%), and lowest for the recording of prescription and procedures [10].

Table 8: Rates of completeness and correctness of diagnoses (data extracted from [4])

<table>
<thead>
<tr>
<th>Condition</th>
<th>Completeness</th>
<th>Correctness</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Cancer¹</td>
<td>45%</td>
<td>57%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diabetes³</td>
<td>53%</td>
<td>89%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diabetes²</td>
<td>18%</td>
<td>97%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diabetes²,³</td>
<td>-</td>
<td>97%</td>
<td>-</td>
<td>100%</td>
</tr>
<tr>
<td>MI¹</td>
<td>37%</td>
<td>46%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Epilepsy¹</td>
<td>-</td>
<td>44%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Epilepsy²</td>
<td>30%</td>
<td>78%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hypertension¹</td>
<td>-</td>
<td>69%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Thyroid diseases¹</td>
<td>-</td>
<td>48%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Thyroid diseases²</td>
<td>42%</td>
<td>84%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asthma²</td>
<td>33%</td>
<td>72%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asthma²,³</td>
<td>-</td>
<td>-</td>
<td>46%</td>
<td>65%</td>
</tr>
<tr>
<td>Glaucoma²,³</td>
<td>-</td>
<td>92%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>CHD²,³</td>
<td>-</td>
<td>-</td>
<td>47%</td>
<td>92%</td>
</tr>
<tr>
<td>Schizophrenia³</td>
<td>-</td>
<td>74%</td>
<td>64% (strict criteria)</td>
<td>89% (broad criteria)</td>
</tr>
<tr>
<td>Ulcerative Colitis⁴</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94%</td>
</tr>
<tr>
<td>Crohn’s disease⁴</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>97%</td>
</tr>
</tbody>
</table>

Gold standards: 1= hospital record, 2= prescription data, 3 = paper notes, 4 = GP’s self audit.
4.3.2. Table 8 gives a summary of detailed results presented in [4]. The authors of [4] carried out a systematic review of several data quality audits over a number of practices. Different practices have different rates of completeness and correctness. Table 8 gives the range of values for completeness and correctness (where given) for the practices in the systematic review. The “Low” column gives the level of completeness/correctness for the practice with the worst audit results and the “High” column gives the level of completeness/correctness for the practice with the best audit results. Where there are results from more than one study for a given condition, multiple rows are presented (diabetes, epilepsy, thyroid conditions and asthma). The gold standard used by each study is denoted by the superscripts. Where no results (or only one result) is recorded a dash appears in both one column.

4.3.3. The difference between the rates in the two columns gives the range of audit results for that condition and measure. Some conditions show high variability (diabetes for example) whereas some show little or no variability (glaucoma). Rates of correctness tend to be higher than rates of completeness. Furthermore, [13] (a study that used probes as a quality measure) found that 96% of diagnosis codes were correctly recorded, in that they were corroborated by other data in the record. Where no diagnosis was recorded in a patient’s record, only 0.5% of patients’ records suggested that a diagnosis was missing.

4.3.4. Whilst interventions have improved the data quality of particular practices, the quality of recording in general has apparently not improved over time. However, as some diagnostic and procedural codes are increasingly also used to drive reimbursement, systematic errors that tend to overestimate the complexity and risk of a patient’s condition have been observed ([9],[14]).

4.3.5. The reported investigations undertaken to date have tended to use a variety of different methods and measures, which make data difficult to compare. For example, studies use a variety of different gold standards (paper records, observation of the clinical encounter) and data quality probes. None of the studies of existing data quality have used prevalence rates. The studies also use a variety of different, and at times somewhat ad hoc, measures with most choosing not to apply the Hogan and Wagner criteria of completeness and correctness. More troubling, there is evidence that multi-practice study designs have tended to omit underperforming practices [4]. For these reasons, the rates of completeness and correctness shown above are indicative only. This, taken with the wide variation in completeness and correctness rates, means that extrapolating from these figures is unlikely to give a meaningful picture of the level of data quality in UK general practice.

4.4. Published studies on data quality improvement

4.4.1. Many studies have addressed data quality in the health care setting. A subset of these looked at ways in which data quality might be improved. A 2006 systematic review of this literature [15] identified 12 studies that i) were empirical studies within general practice, ii) validated the medical record, iii) consisted of an intervention that aimed to improve data quality and iv) that presented at least two measurements of data quality. Six studies were drawn from the UK and one each from the US, Canada, Australia, the Netherlands, Spain and France. Four studies included at least some paper records.

4.4.2. Only three of the 12 studies had control groups, and only one of those fulfilled the basic requirements of an intervention study (randomising to the control group). In only
two of the studies were the interventions described in sufficient detail to enable them to be reproduced. No studies used the Hogan and Wagner (completeness and correctness) criteria (defined in [5]).

4.4.3. Nevertheless, we can extract several lessons from these studies: i) that data quality is highly variable between practices, ii) that few practices have universally good or poor data quality, iii) that data quality does not improve in the absence of some form of intervention (studies from 20 years ago show similar rates of data quality as contemporary studies) and iv) that some types of data (prescriptions, procedures and conditions with clear diagnostic procedures for example) tend to be recorded at a higher level of completeness and correctness than others (observations, nebulous diagnoses for example).

4.4.4. Interventions ranged from minimal (self audit, the instillation of a new GP system) to the involved (group meetings, individual feedback, comparison with the performance in peer practices). The systematic review found that interventions aimed at improving data quality in electronic records were not markedly different to those aimed at improving data quality in paper records. This would seem to indicate that he studies’ designers did not consider the availability of IT systems to be a particular help (or hindrance) to good data quality. All studies showed some improvement in data quality outcomes but the most obvious improvements were due to sustained assessment, feedback and training over periods of two to nine years.

4.5. **Conclusions from the data quality literature**

4.5.1. The literature on data quality in primary health care has addressed several questions, namely: how effective data quality measures are; how good existing primary care data is and what approaches improve data quality.

4.5.2. Studies of the quality of electronic patient data have, to date, been largely quantitative, with the number of items in patient records being the unit of analysis in most cases. Those reviewed by [4] showed that the quality of data is highly variable between practices and between conditions. Of the 12 studies on data quality improvement reviewed by Brouwer, Bindels and van Weert [15], only one was a methodologically sound intervention study, the others being simple before/after study designs. Furthermore, in only two of these studies were the interventions sufficiently well described to permit replication. We found no other intervention studies, and no studies that applied ethnographic methods, to determine the efficacy of data quality improvement measures.

4.5.3. From the body of literature concerning the quality of data in primary care that we have studied, it appears that: i) data quality has not improved over time (as a side effect of the greater prevalence of electronic record systems for example), ii) data quality varies according to the type of data being recorded and iii) data quality can be improved by a systematic program of support and training.

4.5.4. We believe that there is potential for further studies to address the areas that the literature has so far overlooked. Most pressing is the need for studies that address the issue of data quality where the context of use differs from the context of production. More wide ranging intervention studies that address the use of different data quality improvement strategies are needed. Ethnographic studies that address how these interventions work to improve data quality would also be important additions to the literature.
4.6. **Wider perspectives on data quality**

4.6.1. **Findings from the Hampshire pilot**

4.6.1.1. Central Hampshire Electronic Health Record Demonstrator (CHEHRD) was one of 18 NHS Information Authority ERDIP projects running between 1999 and 2003. This national series of pilots was part of the preparation for the widespread adoption of electronic health records. These pilots aimed to identify challenges and test solutions, and to share successful approaches and lessons learned across the NHS.

4.6.1.2. The specific objective of the Central Hampshire Electronic Healthcare Record Demonstrator was to develop an electronic healthcare record which supports 24 hour emergency health and social care for residents in Andover, Mid Hampshire and Eastleigh (collectively, central Hampshire).

4.6.1.3. CHEHRD reviewed informatics and data to be taken into account when constructing its information and technical architecture. A large set of demographic, clinical and social care data items were defined. This set was used as the basis for a common data repository to be shared across the projects' care providers and accessed via new dedicated web applications.

4.6.1.4. Work was undertaken to explore the data quality and clinical governance issues in using GP data as a principal source of clinical content for the shared repository. The clinical governance specification report [36], relating to data quality, is quoted below.

4.6.1.5. “[The NHS Information Strategy] makes the assumption that clinical recording is both accurate and consistent, and evidence from audits of paper records, analysis of GP electronic records, and the Winchester HIS data have all shown that this is rarely the case… Apart from prescribing and laboratory information (which is expected to have a high level of accuracy), work to improve data quality will be required through providing education and feedback to clinical staff. From the experience of the Collecting Health Data from General Practice project, it is likely that at least two cycles of abstracting data, analysis, feedback and education of staff will be required before data of usable quality can be obtained.”

4.6.1.6. A pilot evaluation of the repository's success in supporting emergency care was undertaken in 2003 [37]. Although users only found clinical data of relevance in 20% of cases where it was sought, they considered the system to be both useful and reliable. The authors note that the patient records were often incomplete when compared with standard data sets (e.g. for cancer care) and that there was considerable variation between practices and between patients in the data items collected for the management of chronic diseases. They also point out that a diversity of terminology systems in use across primary and secondary care could complicate the meaningful integration of records, although the demonstrator did not specifically encounter this as Read codes were the only ones in use. The authors identified the following issues as meriting specific attention for the success of a national care record service:

- the implementation of systems that encourage consistent and complete clinical records;
- the training, development, and incentivisation of clinical staff to enable them to record consistently in these systems;
- the development of analysts with a good understanding of the clinical process;
• provision of analytical tools that enable the management of complex episodic data in a reliable and reproducible way.

4.6.1.7. The authors also recommend that the work to assess data quality should be ongoing and open to revision. “Not only can software upgrades change the data structures, but data quality initiatives, training, and changes in staff alter the quantity and quality of information recorded, resulting in the structure and meaning of data changing over time.” Whilst changes in data structures can change the meaning of data, data quality interventions themselves can change the perceived meaning of records over time. For example increasing data quality may change the level of confidence in the electronic record such that absence of a record of a condition is eventually interpreted as absence of that condition in the patient. Thus data quality measures have to change as the expected data quality rises.

4.6.1.8. Certain parallels are apparent between the CHEHRD and the findings of this evaluation. Most notably that: a) accurate and consistent recording is not universal in health care records, b) multiple iterations of assessment, feedback and training are required in order to address these issues c) skilled analysts are required to facilitate this process and d) data quality measures must be reviewed regularly.

4.6.2. Comparing the English SCR and the Scottish ECS

4.6.2.1. The Scottish Emergency Care Record (ECR) is not a direct comparator for the English SCR. The Scottish ECS has a much more clearly defined scope. The record is, and is intended to remain, limited to prescriptions and ongoing conditions, though a linked project is developing the Palliative Care Summary (PCS). These data items are relatively concrete and the end user is able to check whether the two items are mutually consistent. Our own fieldwork observing the ECS in action suggests that end users routinely exercise judgement as to whether the record is likely to be complete or accurate, and act accordingly. Furthermore, the guidance at the point of use for the ECS states that the record should be checked with the patient and their carer.

4.6.2.2. However, the assurance of data quality in the Scottish ECS does not rest solely with the end user. Indeed, it would be highly misleading to view the ECS as an example of a networked electronic record that has no data quality standards. On the contrary, the introduction of the ECS occurred against a background of a long history of data quality initiatives in NHS Scotland, with comprehensive data quality analysis as far back as at least 1994 [38]. A wide ranging process of data quality assurance (DQA) is undertaken by specialist DQA staff. These staff engage with the data producers: “On completion of the assessment DQA staff discuss each clinical coding error (and supporting evidence) with members of the coding staff and agree any actions required,” [39]. The ECS record therefore builds on an existing set of high quality data. Furthermore, the Primary and Community Care Directorate has committed to funding IM&T facilitators on a recurring basis, allowing Health Boards to plan for stable support [40]. As with the CHEHRD pilot, commitment to an ongoing data quality is apparent, despite the existing relatively high base in Scotland.

4.6.3. Standards for local detailed records

4.6.3.1. The term data quality has traditionally focussed on the completeness and correctness of the recording of key clinical facts about individuals. It has also mainly focussed on
fitness for the purpose of aggregating data for research or audit, although in general
practice the use of data for disease registers for individual patient monitoring is an
important driver for improving quality. The primary purpose of the NHS SCR is to
support multi-site and multi-professional shared clinical care. The data quality literature
does not include any analyses of the requirements for fitness for this particular
purpose, and the investigations into data quality reported in this section do not
consider the other properties of data that are need to make it trustworthy, and its uses
safe, for shared care.

4.6.3.2. Safe and well-informed clinical shared care is, however, one of the main drivers for the
adoption of electronic health records internationally, and there is a strong literature on
the requirements for electronic health records. Whilst the EHR is different in essence
to the SCR, the quality of data in the EHR (the LDR in UK General Practice) will
inevitably have an impact on the quality of the data in the SCR, which will be populated
from this more complete record.

4.6.3.3. Major multi-national investigations of EHR requirements include publications by: the
Good European Health Record ([16], [17], [18], [19], [20]), Synapses ([21], [22]) and
the EHCR Support Action projects [23]; academic literature summarised in [24] and
consolidated and internationally agreed requirements published by the International
Organisation for Standardization (ISO) [25].

4.6.3.4. The goal of these requirements is to specify how EHRs should:

- capture faithfully the original meaning intended by the author of a record entry or
  set of entries;
- provide a framework appropriate to the needs of professionals and enterprises to
  analyse and interpret EHRs on an individual or population basis;
- incorporate the necessary medico-legal constructs to support the safely-
  interpretable, relevant and permitted communication of EHR entries between
  professionals working on the same or different sites.

4.6.3.5. At the time of writing this report, ISO is in the process of revising its principal EHR
requirements standard, which is expected to be re-published as ISO [25] in early 2009.
An extract of the some of the new (draft) requirements statements is given below in
Box 2 as an indication of the kinds of wider quality issues that the NHS SCR needs
also to meet, as they contribute to its trustworthiness and/or provide the clinical context
in which individual data item values should be understood.

Box 2: Extracts of requirements from the draft ISO 18308 standard

<table>
<thead>
<tr>
<th>Clinical content requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>The EHR shall be able to represent any kind of health record entry authored by any authorised user, including health professionals of any specialty from primary, secondary, tertiary, community or complementary health care organisations, subjects of care and their representatives.</td>
</tr>
<tr>
<td>The EHR shall be able to represent reported, assessed and measured observations.</td>
</tr>
<tr>
<td>The EHR shall be able to represent opinions, suggestions and hypotheses.</td>
</tr>
<tr>
<td>The EHR shall be able to represent intentions, goals, care plans and actions performed.</td>
</tr>
<tr>
<td>The EHR shall be able to represent concerns, risks, alerts, precautions or warnings about situations to be avoided or activities not in future to be performed.</td>
</tr>
<tr>
<td>The EHR shall be able to represent preventative and wellness information such as health status</td>
</tr>
</tbody>
</table>
assessments, prophylaxis measures and lifestyle.
The EHR shall be able to represent psychological, social, environmental, family, and other life
circumstance information.
The EHR shall be able to represent consents, directives, contracts and mandates for care, and
other legal documents relating to health status and to health care.
The EHR shall be able to represent self-care information, points of view on personal healthcare
issues, levels of satisfaction, expectations and comments authored by the subject of care
and/or authorised representatives and carers.

Clinical information context requirements
The EHR shall preserve the original headings and sub-headings used to organise sets of
record entries.
The EHR shall preserve the original data values within an EHR entry including code systems
and measurement units used at the time the data were originally committed to an EHR system.
The EHR shall represent lists of data items and/or data values within a health record entry such
that their original intended order is preserved.
An EHR entry shall be able to represent links between requested, performed and reported
healthcare activities (e.g. linking a test request to a performed test and to its result).
An EHR entry shall represent any qualification of coded entries by negation, degree of
certainty, severity.
The EHR shall enable an author to explain or justify his or her reasoning or assertions.
An EHR entry shall represent and appropriately label information provided by a third party (e.g.
a family member), another institution (e.g. providing a laboratory result) or a physical device
(such as a cardiac monitor).
An EHR entry shall represent and appropriately label information concerning a third party (e.g.
an item of family history)
The EHR shall be able to represent dates and times imprecisely (to different granularities e.g. a
date as a month and a year or only a year, time as an hour)

Medico-legal requirements
The EHR shall ensure that users who author or authorise entries in a health record are uniquely
and reliably identified.
The EHR shall be able to represent the profession, status and role of any parties identified or
described as care providers within the EHR, and identify an organisation responsible for
sanctioning that role if applicable.
The EHR shall represent the care setting, organisation and physical location at which a
recorded health care activity has occurred.
The EHR shall represent the date and time at which each health record entry was originally
committed to a health record.
The EHR shall enable an authorised user to modify a record entry, but not to alter or erase an
original entry.
The EHR shall be able to represent the rationale for revising an EHR entry or set of entries.
The EHR shall be able to represent the consent and policies specified by a subject of care
and/or a representative for the disclosure of his or her EHR.
The EHR shall support the maintenance of an audit trail of the creation of, modification of, and
access to health record entries.

4.6.3.6. The former GP systems accreditation standard RFA (in particular RFA 99) included
tests of many of the properties documented in Box 2, and it therefore to be hoped that
the source clinical data being extracted for the SCR already meets most of these
requirements. However, this evaluation has not been scoped to include an
examination of the data extraction methods implemented by GP systems vendors, nor
of the upload interface, nor of the SCR repository architecture. It is therefore not
possible to state if these EHR requirements are met. Given that the NHS SCR is the seedling of an EHR (as defined by ISO TR 20514 [26]), it is recommended that a more technical / informatics evaluation is performed to verify that the data in it meets internationally agreed standards for EHR medico-legal and clinical data integrity.

4.7. Social factors in technology change

4.7.1. Introducing new technologies into a complex system requires extensive changes in individual roles, relationships, and business processes – the so-called ‘socio-technical’ aspects of change ([6],[27],[28]). Individuals’ resistance to new technologies is rarely due entirely (or even predominantly) to lack of training; it can often be tracked back to conflicting values, priorities, or a mismatch of personal identity with the role changes associated with the new technology (“I’ve been reduced to a data entry clerk”). The routine associated with a new technology must ‘fit’ with other routines within and beyond the organisation. It must be resourced (in terms of funding, physical space, staff, and expertise).

4.7.2. Complex technologies that impact centrally on multiple routines are almost never implemented smoothly; there is an ‘embedding’ phase in which staff identities, work patterns, and interactions gradually change, and various routines are adjusted to accommodate the new technology. This phase typically requires sensitive facilitation and flexibility to ensure that (a) the meaning and vision for the new technology is negotiated (as far as possible) to everyone’s satisfaction; (b) staff voice their concerns rather than bottling up resentment; (c) these concerns are addressed – for example via bespoke training, adjustments to the technology, adjustments to the routine, or provision of additional resources; and (d) competing routines are adjusted appropriately rather than marginalised.

4.7.3. An important dimension of data quality work is the development of effective routines. An organisational routine (as opposed to the lay use of the term meaning ‘personal habit’) has been defined as “a repetitive, recognizable pattern of interdependent actions, involving multiple actors” [29]. Achieving high quality data requires a number of different routines (for example, the routine associated with receiving, coding and scanning hospital discharge and outpatient clinic letters; the routine for checking the electronic record of a patient registering with the practice, whose record has been transferred via GP2GP; the routine for feeding the key findings from a CHART audit into organisational learning). If these routines are weakly developed (for example, if there is no clear line of responsibility for key components, or if roles and handovers are unclear), data quality is likely to suffer.

4.7.4. There is a need to recognise the impact of social factors on technology adoption throughout the SCR programme [41]. The social aspects of data quality are, as yet, poorly understood. Whilst standards aim to drive up quality they cannot, in and of themselves, improve processes. As the next section will show, the skills and mindfulness required to improve routines are often beyond the reach of practices. A skilled data quality facilitator not only ensures that individuals within the practice have the relevant skills and motivation, but also that the routine is well coordinated across different staff.
5. Analysis of the primary research data

5.1. What do experts and practitioners in the field think – and why?

5.1.1. We sought input from two main groups: data quality experts, and clinicians and supporting staff. The clinicians and supporting staff can further be broken down into two groups: data producers (GPs, practice managers, practice IT leads etc.) and data users (out of hours and acute clinical staff, community nurses etc.) For each of these groups we have selected participants from the UK and overseas.

5.1.2. We ran two online focus groups, one consisting of GPs and Practice managers and one consisting of Data Quality experts. We also completed a two round online policy Delphi process. We also carried out a series of interviews. Our interviews took a variety of formats, including site visits, telephone interviews and interviews at UCL. Overall we obtained input from:

- A&E Consultants
- GPs
- Out of Hours and Walk in Centre Nurses
- Nurses from NHS Direct
- Data Quality experts within CfH and PRIMIS
- International experts on clinical data standards
- ~15 personnel from early adopter practices, including:
  - IT Leads
  - Practice Managers
  - GPs
  - Nurses (Practice and Walk in Centre)
  - Summarising personnel

5.2. The perceived importance of data quality

5.2.1. In the focus groups, Delphi process and interviews, data quality at the point of recording clinical information was generally perceived as being important. A range of reasons were given for this. In the focus groups, three GPs independently mentioned that good data quality permitted them to make informed decisions about the population that they served. They stated that good quality data helps to answer the question “how good a doctor am I for this population” and gives “a good picture of local care”. Another GP uses this population data to lobby for extra resources to support local initiatives that the data suggested were necessary.

5.2.2. Not all GPs were motivated by intrinsic factors such as these: in these same focus groups, two GPs mentioned QoFs as being their driver towards good data quality, one other mentioned the need to “standardise” patient records in order to prove that the practice was performing as well as its neighbours.
5.2.3. Interestingly, non-clinicians were the only ones who mentioned the patient’s right to have good quality data in their record. In the course of an interview, one data quality facilitator stated that the overriding motivator for good data quality was that the “patient expects their record to be right”. No clinicians made equivalent statements in their focus groups or in their interviews – however it may be that clinicians believe that this is so self evident as not to require stating.

5.2.4. Not all respondents were unequivocal in their support of the drive for good data quality. Several respondents were eager to stress that data quality is one of many competing requirements. For example, a practice manager asserted: “we must not be slaves to data quality for its own sake” in the practitioners and users focus group. Some went further, maintaining that the premise of reducing important information about the patient to data items that are easily validated is a mistake: “Good medicine is not about these hard facts recorded in SCRs”.

5.2.5. Overall, most clinicians and practice managers saw data quality as important, either for its own sake or because it is part of a requirement of the QoF or the IM&T DES.

5.2.6. Positions on the importance of data quality at the point of use were similar across the board. The Delphi exercise asked respondents whether a drop in data quality would affect the degree to which they felt the SCR would be useful. When presented with scenarios in which data would be 90% complete and 98% correct versus scenarios in which data would be 70% complete and 90% correct, few respondents changed their priorities regarding the appropriate clinical uses they would make of the SCR. None of the clinical respondents changed their priorities.

5.2.7. However, many respondents suggested that they would give less weight to the SCR and more weight to other sources of information about the patient: “I would need some other way of checking the data” (Mental Health Lead, Delphi questionnaire); “The need for examination of the patient increases” (French expert on health informatics). In the main, the position of the clinicians could be summed up by the following contribution to the Delphi from this same expert: “Any chance to improve information [by accessing the SCR] must be grasped”.

5.3. **Good data quality does not just happen**

5.3.1. This section will highlight some of the data quality issues that have been identified during our primary research. In their actions so far and in commissioning this evaluation it is clear that CfH are convinced of the need to address data quality. Nevertheless, we will present here a re-statement of some of the problems that we have identified with DQ for two reasons: firstly, it helps to circumscribe the extent of the DQ challenge; secondly it provides support for the business case for investment in DQ (as supported by the results of this evaluation).

5.3.2. Mistakes can be classified in several ways. One classification is by type of error. Firstly, data can be missing, out of date or factually incorrect. Secondly, data can be correct but incorrectly coded. An example of the first mistake was reported in one of the focus group discussions:

"A summarising team found that a note in a woman's clinical record saying that she had a heart condition was actually for another person. Over a period of 8 years the woman had visited various hospital departments - each time the GP's summary and
the consultant discharge letter noted this heart condition, nobody questioned it and she was treated accordingly. After 8 years the GP then had to explain the mistake to her." (Comment in the CfH focus group).

5.3.3. An example of the second type was also reported in one of the focus group discussions:

"Some of the findings included patients being prescribed expensive drugs without a corresponding diagnosis; more patients recorded as having diabetic retinopathy than diabetes alone; and women who appeared to have had several hysterectomies." (Participant in DQ Expert Focus group).

5.3.4. Another way of classifying data quality problems is by the source of error. Some DQ issues arise from systematic mistakes. For example

"A search of the clinical system revealed that the entry form used to monitor diabetics was using the diagnosis code for diabetic retinopathy rather than the screening code." (Participant in DQ Expert Focus group).

5.3.5. Mistakes can also arise in an ad-hoc way, due to human error, or the difficulty of dealing with unusual circumstances. For example, an audit in one practice found data entry problems arising when data came from a non-standard source (in this case a solicitor’s letter) for which there were no existing routines.

5.3.6. Data can be damaged maliciously:

"This practice was going through a "re-organisation", and one very aggrieved member of staff decided that they would delete patient records that were held on the computer system. She did this while logged on as an admin user - using the same details that were used by the other eight admin members of staff." (Participant in DQ Expert Focus group).

5.3.7. Finally, data in the SCR may be a faithful representation of the patient's circumstances, but may still be misleading because it is being used in a different context to the one in which it was recorded. For example, during a field visit to a walk in centre in one of the early adopter PCTs a patient presented with a badly cut leg with persistent bleeding. The SCR confirmed the drugs that the patient had listed. In addition the SCR reported that the patient was taking aspirin daily as an anticoagulant. The nurse did not consider this relevant to the patients treatment. However the patient could have been advised to cease taking the aspirin until the wound had healed. The SCR was seen as a source of possible contraindications in this context, rather than as a source of data for clinical decision making. Whilst the error is, in the strictest sense, one of data use rather than data entry, the framing of the data on the SCR may have been responsible for the nurse overlooking this data.

5.4. **Practices need training and support**

5.4.1. Among the clinicians and experts that we contacted, both in the UK and abroad, there was clear recognition of the fact that the move to EHRs in general, and to the SCR in particular, required some new skills. However, the need for training was perceived as going beyond IT training, and in fact the IT skills were rarely mentioned. Rather, the respondents to our research identified a set of other training needs that they judged to be more important than the need for IT skills.
5.4.2. This view of training is summed up by the following quote from a UK clinical data quality expert in one of the focus groups:

“Practices need two sorts of training, the how ... and the why. This latter is the most important as the system users learn from the understanding they gain and can extrapolate that information to new scenarios.” (Emphasis ours.)

5.4.3. This view, that the data producers (the GPs) must understand what they are doing and why, rather than simply being competent to use the technology delivering the electronic patient record, was echoed elsewhere. For example, in one focus group, a DQF expressed concern that GPs “lack understanding of information management... The clinical side is second to none but there is often a rather vaguer approach to data entry”. In another focus group, a GP addressed this same issue: “We need to encourage colleagues to only record confirmed diagnoses [and not the] assumptions that are made”.

5.4.4. The vast majority of clinical staff have the procedural knowledge necessary to use IT systems, with varying degrees of enthusiasm and competence, and there is still a need for some IT skills training. However, far fewer clinicians have been helped to gain the knowledge of the range of uses to which the recorded data may be put, or the importance of understanding the way in which the data will be represented and processed within a clinical system. This personal knowledge is much harder to imbue, and the end product is not a set of skills per se. Specifically, to be able to exercise good judgement when recording data, GPs need to be aware of why they are recording data, and this awareness can only come if they understand the wider context of how the data will be stored and used. This does not have to be explicit knowledge however. The development of good routines driven by the tacit awareness of what constitutes good data is sufficient to drive up quality. For example, an experienced DQF told the focus group how:

“The IM&T DES...raises awareness of processes, and how individual roles contribute towards the process, rather than just looking at the data and making individual corrections. [Once an audit has identified some errors] GPs were then very willing to look at how they were recording this. The would also start looking then at what else they may be missing.” (Emphasis ours.)

This illustrates how practices can become more mindful of the ways in which they should be recording data without necessarily becoming experts in the end use and the data architecture. However, they require the support of someone who is expert in those areas to guide them in developing that tacit awareness.

5.4.5. It is unsurprising that data quality experts are aware of the need for mindful practice. From the Delphi exercise it was apparent that several of the clinical respondents also identified something like mindfulness as being an important factor in how good a practice’s data quality would be. One respondent said that the worse practices probably “do not understand the implications of the SCR not being accurate or complete”. Another respondent suggested that the better performing practice had “awareness of the benefits of an accurate data set and good patient records.” Interestingly, both of these clinicians were on the “using end” of the SCR, and so presumably were mindful of these issues themselves.

5.4.6. However, this mindfulness will only come about if a practice is committed to investing their time and energy to this rather than other, competing requirements. One data quality facilitator contributed the following to the focus group: “It depends on the ‘will’ of the team involved... we can give them as much information as we like, give them as
much training, even lots of money but if [the practice manager and at least one GP] don’t work together... then we are facing a very, very difficult challenge.” In the Delphi questionnaire, another data quality facilitator explained how a practice-wide attitude could affect the access to training among staff. She told us that she had worked with practices where the nurses “were not confident using the computer and the GPs did not encourage them to attend IT training.”

5.4.7. IT skills and mindfulness are both essential if practices are to produce good quality data. Nevertheless, the training needs of practices differ. For example, in one practice site visit we met with a very skilled IT lead. However she was struggling with several of the CHART queries because she was not sure what they were “for”. Whilst she could (and did) execute them nothing was done with these results, because no one had explained what it was they were indicating. A practice manager that we interviewed also confided that she hated CHART queries and felt that they “created a lot of unnecessary work”. She too had not had support in interpreting and applying the implications arising from these queries and so did not understand their relevance to her practice management role.

5.4.8. The data quality facilitators that we interviewed felt strongly that tailored training was necessary. One told us that it was her experience that some practices needed more help than others, especially the ones that have developed their own routines but that have “made the same mistakes for years.” She found that visiting practices in situ and observing their processes was much more effective than inviting practices to a workshop. Another DQF had found that participants in workshops were often frustrated because they lacked the background knowledge to benefit from the training in that environment: “The information was difficult for practices to understand as they weren’t used to this kind of work.”

5.5. The data quality facilitator is a specialist role

5.5.1. PRIMIS have authored a data quality facilitator job description that states inter alia that DQFs must “Organise feedback sessions at each practice to consider data quality issues arising from the analysed data (in liaison with PRIMIS Team).” This task covers a multitude of activities which vary from practice to practice. Where the data quality facilitators work best, they encourage the mindfulness that practices need in order to increase their data quality.

5.5.2. Data quality facilitators involve themselves closely with the practices. A data quality facilitator described the interventions that she has carried out. As well as prompting GPs to look at how and why they are recording data, she also helps practices draw up action plans for improvement. Action plans should be “wide ranging and practice specific [and] developed with the practice IT lead,” in order to be effective.

5.5.3. There are, of course, occasions when this does not go according to plan. For example, one skilled IT manager told us of her discomfort that she doesn’t get feedback on the results of the CHART audits that she performs. She does not feel able to act on these without the support of a DQF. The manager of a different practice with ongoing data quality problems also complained of being in the same position. It is telling that, in both of these cases, contact with data quality facilitators has been limited. For example, in the case of the practice manager: “they [the PCT] initially sent someone around to explain what needed doing [and] left me a list to work on, but nothing else since then.”
5.5.4. Where the data quality facilitators are well integrated they are well regarded. The characteristics that interviewees mentioned as being important in a DQF were: flexibility; willingness to “get their hands dirty”; being good at explaining and being patient. The data quality facilitators we spoke to bore this out. Interviewees made comments such as:

“[We need] communication skills, speaking to people at all levels, not in a techie way but understanding the practices.”

“[We have to understand] what other things the practices are trying to do sat the same time, [their] conflicting priorities.”

“We must be able to do anything and everything… to be prepared to wait [and] to treat practices individually.”

5.5.5. These skills take time to develop, and training data quality facilitators is not straightforward. One of our interviewees (an experienced data quality facilitator) told us that DQFs were trained in a very broad way and the most important aspect of their knowledge is a tacit understanding of the kinds of support different practices need. She told us that “you can never come up with a proper handbook and say ‘this is it’. Facilitators build their skills through their experiences.”

5.6. The impact of existing data quality standards

5.6.1. There is a wider DQ agenda in UK primary care. At the time that the participants in the Early Adopter phase of the SCR were being selected there were several data quality initiatives. These included: the Quality and Outcomes Framework (QOF); the Primary Care Information Service (PRIMIS) tool “Care and Health Analysis in Real Time” (CHART); the Paperlight accreditation process and the Information Management and Technology Directly Enhanced Service (IM&T DES) programme.

5.6.2. All PCTs studied appeared to be well-managed organisations with strong leadership and a tradition of ICT innovation. They made good progress in the set-up period, though this was hard work all round. Much preparatory work was undertaken by PCTs and GP practices to improve data quality for the SCR, and this linked with wider national incentives on primary care data quality. Whilst many participating practices began with mediocre data, the process for data quality improvement was good, and rapid progress was made in many (though not all) practices. Participation in the SCR programme appeared to be a strong incentive for data quality improvement initiatives that had wider relevance to PCTs and GP practices.

5.6.3. Data quality practice as it stands is excellent in places, whilst in others there is cause for concern. We must stress here that the practices that we consider to be the best performers are not necessarily those with the best results in the CHART queries. Instead we consider a mindful approach and well thought out processes to be more appropriate hallmarks of a practice that is performing well.

5.6.4. The reason for valuing process oriented outcomes, rather than the quantitative, audit-based measures, is that there are practices that we have encountered that are currently taking the approach of “fixing” individual records that the CHART queries
identify as problematic. This does not amount to deliberate gaming on the part of these practices as they genuinely believe that the purpose of the CHART queries is to prompt this course of action. Therefore there are several practices among the early adopters that will appear to have good data quality according to the CHART audit, but where the underlying data quality problems have not been identified.

5.6.5. Among many of the early adopter practices we have studied the data quality process is driven by the CHART queries. In most early adopter practices these queries are run in order to identify problems, defined by one practice IT lead as the process of determining “who’s not doing something.” If the audit shows that there is a problem with the data, the data quality facilitator and IT lead look at the practice’s data production process. There may be other data entry errors that the CHART queries are unable to pick up, but that arise in the same way as the ones identified. One example problem may identify an entire problematic pathway.

5.6.6. In the best practices protocols have been developed to support the production of good data. IT leads have described processes in which they forged collaborations within the practice to design these protocols. One IT lead described the process as “thinking how would this best work with the way we work.” The results of the CHART audits, together with the discussion of the practice’s way of working, continue to lead to the development of new protocols.

5.6.7. Not all practices work as well as this. Some are still using the pro-forma templates that they had been given in PCT training courses and had made no attempt to tailor these to their own requirements. Others still see CHART as being indicative of records that must be fixed, rather than of processes that need to be altered. These practices give cause for concern, since to an external clinical assessor they may appear to have good processes in place. For this reason we believe the data quality facilitator is essential in assessing a practice’s readiness to upload, given their close involvement with, and better insight into, the workings of the practice.

5.7. The effectiveness of current measures.

5.7.1. We have found many instances in which the current data quality standards work extremely well. However, in order to understand why they work so effectively, it is important to examine what those standards consist of.

5.7.2. In most early adopter practices the measures in place consist of a process of audit driven improvement supported by a data quality facilitator. In order to assist the practice in assessing its performance, the DQFs use the CHART audits to provide a benchmark. In general, there are two types of benchmark: absolute and comparative. An absolute benchmark consists of a measure (or measures) and a threshold. An individual that fails to reach that threshold effectively “fails” the benchmark. A comparator benchmark consists of a set of measures, a cohort to which that measure will be applied, and a level of tolerance that will be applied. Absolute benchmarks are appropriate where known levels of performance (such as minimum safety standards) are required, and can be quantified. Comparative benchmarks on the other hand address the situation where known standards cannot be determined a priori. Comparative benchmarks do, however, indicate that further investigation is required where a practice’s performance deviates widely from that of its peers.

---

\(^2\) We here define gaming as adherence to the letter, rather than the spirit, of a measure. This is usually the case where the participant knows that adherence to the measure will be rewarded, but that innovative ways of reaching the same ends will not be.
5.7.3. The DQFs support practices in comparing their data with a comparator benchmark. This benchmark consists of the audit data from a cohort of practices in the same PCT. There are plans in place to permit the comparison between practices of similar types also. Practices are encouraged to reflect on the reasons why their performance may differ from that of their peers. This has several advantages: firstly, the benchmark compares performance against a cohort of similar practice and should therefore be reasonable and attainable. Secondly, our data have shown that staff in practices are motivated by the desire to perform as well as, or better than, their peers. Finally, comparative benchmarks provide a “self-setting threshold”: the performance of the best practices is the standard against which all practices are compared. This is a more appropriate standard in this situation, where absolute data quality standards cannot be fixed in advance.

5.7.4. In the best cases the PRIMIS CHART audits are well understood, and used in a diagnostic capacity. More important is the ongoing commitment to reviewing and improving data production routines in the light of the findings from these audits. In the most successful practices, this process has been supported by data quality facilitators. In many cases these have been provided by the PCT. In others the practice has hired its own experts. Furthermore, in these best cases, the practice as a whole is involved in the drawing up of data production routines.

5.7.5. However, in many practices, even among the early adopters, one or more of these factors is not in place and data quality (as measured by the CHART audits) and staff morale (as evidenced by the degree of commitment to data quality) have suffered.

5.7.6. We have found practices in which the purpose of data quality interventions was not understood. In one practice we examined, the Practice Manager volunteered, “I hate CHARTS…always guilty of creating a lot of extra unnecessary work… Partners spent about 4 hours each going through the lists I gave them”. This practice manager has not understood that the purpose of CHART is to identify the source of problems, rather than to fix them as they arise. Until this practice takes steps to improve data production routines, the CHART audits will continue to identify a lot of (necessary) work. In other practices, genuine efforts were being made to improve data quality, but the lessons from the CHART data were not being learnt. Another practice manager was given a number of indirect prompts to elicit whether the CHART results led to changes in template protocol, provided by the PCT. Instead there were many references to having the doctors “fix” the data and none to changing the protocol as a result of audits. They did however say that the protocol would be reviewed annually, as this is a QOF requirement. The chart results were not considered to be an input to this process.

5.7.7. We also found practices in which the CHART audits were, in the main, well understood, but that input from a DQF had ceased. For example, we studied one practice with a specialist IT manager. The PCT had requested that the CHART queries, which the IT manager referred to as the IM&T DES queries, were not intended as a data quality measure. This confusion is understandable as to the IT manager CHART is a data quality tool. She expressed her dissatisfaction that she had received no feedback from the PCT in response to these queries, and so did not know how “good” her data was. A skilled DQF would have identified this source of confusion and clarified the situation with the IT manager.

5.7.8. These examples demonstrate that the absence of one of the four measures (audits, attention to routines, support from a DQF, commitment across the board) can seriously undermine the ability to generate high quality data.
5.8. **Clinical users will exercise judgment**

5.8.1. Data quality is important, but few clinical records (paper or electronic) are 100% complete or accurate, and it is an established part of clinical practice to work on the assumption that data may be incomplete or incorrect, regardless of source. For example, just because a patient has no recorded allergies doesn’t mean they have no allergies. If the data held in the SCR are known, or generally perceived, to be of poor quality, then it is likely that clinicians will choose not to consult the SCR. Thus the risks associated with poor quality data are not so much that clinical error will increase (although that is a real possibility), but that the SCR might fail to deliver sufficient benefit to users because clinicians fail to trust or use it.

5.8.2. Clinicians were prepared to exercise judgment in deciding whether to use the SCR. The Delphi exercise asked respondents to state how they thought the SCR should be used, given two hypothetical levels of quality. The higher level was described as an SCR that would be 90% complete and 98% accurate and the lower level was described as an SCR that would be 70% complete and 90% accurate. For the majority of respondents, their priorities did not change. As one respondent, a former GP, succinctly stated that "Information is information". This was echoed by another respondent, an international Health Informatics expert: "Whatever the level of completeness... any chance to improve information is to be grasped". However, several mentioned that they would expect the information in the SCR to be treated with greater caution at the lower levels of completeness and accuracy ("No changes in priority. However, I would need to have some other way of checking the accuracy of the data.", "I think that the 'weight' within which the clinician can use the information to aid their decision will be dramatically reduced.")

5.8.3. Six out of 15 respondents stated that they would seek corroboration of the SCR data regardless of the level of data quality. A further four from 15 stated that their use priorities of use of the SCR wouldn’t change, but that they would seek extra validation of the data at the lower levels of accuracy and reliability, and four from 15 stated that they would change their priorities of use at the lower levels of accuracy and reliability.
6. Conclusions

In this section we will address the three questions initially posed by the funder, to wit:

i. “If no data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records occurred, what would be the risks in terms of the completeness and accuracy of the derived summaries?”

ii. “If no data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records occurred, would this have implications for the subsequent upload of problems and diagnoses?”

iii. “If it was considered desirable that minimum data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records, what might those data quality standards be in order to maximise patient benefit and minimise patient risk?”

6.1. Initial risks in the absence of data quality standards.

“If no data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records occurred, what would be the risks in terms of the completeness and accuracy of the derived summaries?”

6.1.1. A systematic review of 24 studies of the prevalence of errors in morbidity coding in UK general practice computerised records [4] found completeness (sensitivity) of records compared to the gold standard to range between ~30% to ~100%. Correctness (PPV) of these records ranged from ~60% to ~99%. This accords with a study of Dutch community pharmacy records [30]. In this study, average completeness values ranged from 29% (for heart failure) to 84% (for diabetes). PPV ranged from ~20% (Angina and Asthma) to ~100% (Cardiovascular diseases and hypothyroidism). On the whole [4] found that the clearer the diagnostic features of the condition in question, the more likely it was to be correctly recorded.

6.1.2. The review carried out by [4] also showed that any improvement of data quality over time was difficult to ascertain, and that several studies had discarded practices that were unable to provide data – thus biasing the sample towards better recorders of data.

6.1.3. Extrapolating from these results, it is likely that completeness and accuracy would be variable, with some morbidities being better recorded than others, and some practices being better than others. In addition, [4] showed that data quality improvement does not come about in the absence of an intervention programme.

6.2. Ongoing risks in the absence of data quality standards

“If no data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records occurred, would this have implications for the subsequent upload of problems and diagnoses?”
6.2.1. The first data items to be recorded in the SCR, current medications, allergies and adverse reactions have several advantages. Firstly, these are relatively well-defined data items. A prescription is usually unequivocally recorded, whereas diagnoses and problems are much more nebulous (especially in primary care). Secondly, and allied to the first point, research has shown that these data items tend to be the best recorded, with the highest rates for completeness and accuracy.

6.2.2. As the SCR grows, the included data items will become more nebulous and complex. As a result they are likely to be less easily recordable, and much less easy to validate through probes. Consequently, the routines of data production will increasingly become the method by which data quality must be assured for subsequent uploads.

6.2.3. The original question asked by the funders focused on clinical risk. However, the Delphi research that we carried out indicated that clinicians are confident in their professional judgement as to the weight of evidence that can be accorded to any data item. More research is needed to determine the extent of real clinical risk associated with omitted or incorrect data items in the SCR when it is consulted prior to clinical care decisions. If the data quality in the SCR is perceived as being generally questionable, an additional risk is that clinicians will “refuse to use” the SCR.

6.3. Potential data quality standards: option appraisal

“If it was considered desirable that minimum data quality standards were applied to a practice’s clinical database before uploading of NHS Summary Care Records, what might those data quality standards be in order to maximise patient benefit and minimise patient risk?”

6.3.1. In the absence of a clearly delineated set of use cases for the SCR, and a priority order over these use cases, it is inherently impossible to determine a single set of prescriptive data quality standards. Quantitative standards are of limited use as they only capture data quality at one moment in time and they are only applicable to the minority of clinical data. If the potential uses of the SCR are to remain flexible and open ended, specifying prescriptive data quality standards will continue to be problematic.

6.3.2. However, even in the absence of prescriptive data quality standards, it is possible to make reasonable statements about ‘better’ or ‘worse’ data quality. If good data are defined as those which are ‘fit for purpose’, for example, then a working definition of poor quality data might be ‘data that could mislead a clinician in one or more contexts of use’, and a working definition of good quality data might be ‘data that lead to better decision making in one or more contexts of use’. Note that these definitions, whilst pragmatic and defensible even in the absence of specific use cases, do not lead directly to quantifiable measures of data quality.

6.3.3. The approaches to data quality standards that have been tried in other projects that we have identified thus far have been:

1. The withdrawal of standards
2. Auditing plus payment
3. Auditing plus process improvement facilitation
These approaches are summarised in Table 9

Table 9: Possible approaches to data quality assurance

<table>
<thead>
<tr>
<th>Option</th>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Withdraw standards</strong></td>
<td>Very low investment needed.</td>
<td>Data quality is unknown at best and may be unsafe.</td>
</tr>
<tr>
<td><strong>Audit plus payment</strong></td>
<td>Easily managed once the metrics are in place.</td>
<td>Metrics are difficult to design.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Only validates the practice records at a given point in time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In practice this is likely to lead to the fixing of individual records, rather than a commitment to data quality overall.</td>
</tr>
<tr>
<td><strong>Audit plus facilitation</strong></td>
<td>Changes the data production culture, making DQ part of the professional role.</td>
<td>Requires ongoing investment in the skills and professionalism of the facilitators.</td>
</tr>
<tr>
<td></td>
<td>Avoids the unintended consequences of metric driven behaviour.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Is the only practical way to ensure that data is produced in a way that promotes quality overall, in an ongoing fashion.</td>
<td></td>
</tr>
</tbody>
</table>

**Option 1: Withdrawal of standards**

6.3.4. The literature showed that data quality tends to be highly variable between practices and between morbidities. Rates of omitted data items run from 70% to less than 1% and rates of recorded errors run from 80% to zero.

6.3.5. There is no technical barrier that would prevent CfH suggesting that no data quality standards be applied. This is likely be the option with the lowest direct costs, although an analysis of indirect costs (such as medico-legal risk) should be undertaken if a thorough analysis of this option is to be carried out. However, whilst this is the approach that requires the fewest resources, the literature shows that data quality where no intervention has been carried out is highly variable and, in the majority of cases, very low. As our research has shown, clinicians will only use the SCR if they feel it will make a safe and meaningful contribution to their clinical decision-making. The laissez-faire approach will lead to unknown levels of data quality at best, and data quality levels that are so poor as to make the data unfit for use at worst. This will have repercussions for the extent to which the SCR is used in practice. Consequently we do not recommend this approach.

**Option 2: Audit plus payment**

6.3.6. Academic literature on data quality in the primary care sector and other health care settings almost exclusively focuses on probes that can be undertaken to validate data. Probes consist of queries executed on an organisation’s data with a view to identifying data that are inconsistent with expectations. Many studies have been carried out in that consider audit either as an end in itself (e.g. [31],[4],[32]) or as a means of
supporting interventions to encourage clinicians to improve data quality ([9],[10],[12],[15]). This is reflected in UK General Practice, with the widespread adoption of PRIMIS’ CHART queries as an auditing tool.

6.3.7. There are three notable difficulties with the audit-based approach: 1) it is based on the use of proxies for the underlying clinical reality, 2) these proxies are poorly used, as the formal notion of a quality measure is not widely understood and 3) audits are taken at a point in time and therefore reflect data quality only at that particular instant.

6.3.8. To expand on the first point: as reported in the literature, audits use several methods to validate the data, including comparison with prescription data, comparison with prevalence data, re-abstraction of medical records [12] and the use of queries (probes) based on medical rules [15]. Whilst informative, these features are only proxies of the underlying clinical reality. Certain aspects of the clinical record are not susceptible to validation by these sorts of probes: for example where no clear drug to diagnosis relationship exists. Therefore these probes can never be more than a diagnostic tool that identifies potential problems with data production – they can never validate the electronic patient data as a whole.

6.3.9. Furthermore, we cannot be sure how some of these proxies may be applied. For example: prevalence data is meant to give an indication of whether a practice is recording morbidities by comparing the practice rates with external rates. However, as [12] showed, when prevalence rates for morbidities in practice records were measured against the MSGP4 rates, reported rates of these morbidities increased until they not only approached, but actually overshot, the recorded mean morbidity rate. Site visits and interviews have shown that practice managers and GPs do not universally understand what the CHART queries are showing them.

6.3.10. Finally – data quality probes are only valid at the time at which they are applied: records may be complete and correct at the point in time at which the probe is applied. If underlying data production practices are weak however, data quality is likely to deteriorate as soon as new data items are added.

6.3.11. We would also suggest that, as the uses of the SCR record become more wide ranging, devising QoF measures that accurately reflect the requirements for that data will become more difficult. A QoF metric that captures whether the appropriate data have been collected for patients with certain morbidities may be easy to devise. However, devising a QoF requirement such that all patients with a particular condition have a detailed care management plan is much more difficult.

6.3.12. Several observations from the focus groups, the Delphi exercise and the interviews support the conclusion the QoF is a major driver towards coding clinical data. However these same sources have also admitted what is widely known about payment-related quality metrics: data completeness efforts are often focussed on those items that have a business case, and such metrics are in no way a reflection of the generic data quality standards or data quality assurance practices of the organisation or team.

6.3.13. In summary, making the QoF the main driver of data quality requires only limited resources and is likely to lead to high levels of compliance with those audits. However, the QoF exercises are susceptible to a degree of gaming, devising appropriate QoF metrics will be challenging, and the QoF is not a continuous process, which means that it can only ever assure data quality at the time that it is administered.
Option 3: Audits plus facilitation.

6.3.14. Audit measures are only possible for a limited subset of data items. For most clinical data, no meaningful rule based queries can be constructed. In addition, because the current plan is for the SCR to be used in a variety of settings, it is not possible to determine a priori which data need to be audited. Finally, audits such as the CHART queries only reveal the state of one particular set of data at one particular point in time.

6.3.15. These audits do, however, point to underlying problems with the data capture process. Once these problems have been identified, it is then possible to determine the source of the problems (e.g. problems with the process of data entry or coding) which can then be addressed by process improvement interventions: for example, drawing up protocols for data capture or designing and targeting staff training courses. GP practices do not necessarily have the skills to carry out process improvement interventions. Even where practices have staff that are skilled in both IT and data quality, they may need support in carrying out CHART and other audits, analysing them, and identifying which of the practice’s processes and procedures to target for improvement. To this end, the role of the data quality facilitator (DQF) has been developed.

6.3.16. The PRIMIS job facilitator job description states that the role of the DQF is to: “provide information management skills, data feedback, support and training to Primary Care General Practitioners and Primary Care staff in the local Primary Care Trust area”. In practice, DQFs need to understand general practice in particular and primary care in general. They also need to be skilled in IT training, and to be possessed of excellent communication and change management skills. DQFs can suggest or refine a practice’s data quality policy and process, and encourage compliance with those policies and processes. Where this functions well, practices have used the audits as successful triggers to changing their data capture behaviour. In some cases, due to a mismatch between the facilitator and the practice for example, we have identified instances where lessons have not been learnt by practices. This may be because they do not grasp the importance of the CHART queries, or they do not realise that it is insufficient simply to “fix” the records that are flagged as problematic, rather than addressing the underlying source of the error. It is therefore imperative that interventions are carried out by DQFs that have the skills and flexibility to fully engage primary care staff in order to help them change their practices.

6.4. HealthSpace as a marginal source of data validation

6.4.1. In several forums, participants have suggested a role for HealthSpace as an additional data validation. For example Cayton 2006 [33] suggested that: “the public information programme should inform patients that they have a defined and realistic period of time to review their proposed summary record, for example by viewing their proposed summary on HealthSpace or by asking to see a printed copy provided by their GP, should they wish to do so. Patients would be invited to correct or amend their record and offer explicit consent for their record to be shared”. However, we do not expect HealthSpace to make a more than a marginal contribution to data quality.

6.4.2. Pyper et al [34] studied 100 patients who were interviewed whilst using a computer to access their medical record. These users identified a number of errors in their medical details. These are summarised in Box 3.
Box 3: Errors found in medical records by service users (from Pyper et al [34])

- Missing information: medication, vaccinations, allergies, test results, and patient records from before the practice became electronic
- Missing consultations: nurse, health visitor, out-of-hours doctor
- Missing events: adverse reaction to medication, breast screening, operations, tuberculosis, childbirth, premature childbirth, miscarriage, sterilisation, irritable bowel syndrome, severe migraine, glaucoma, fracture, repeated episodes, and minor surgery
- Missing referrals: cardiology, urology, endoscopy, orthopaedic, physiotherapy

6.4.3. However, the subjects in the study needed considerable support to be able to provide feedback on their records. 42% required help in understanding medical terminology, 13% needed help in understanding acronyms and 17% needed help to understand test results. Patient's understanding and recall is a far from authoritative source of validation. Pyper et al report a number of patient misunderstandings and knowledge gaps, which are summarised in Box 4:

Box 4: Misunderstandings and knowledge gaps in service users (from Pyper et al [34])

- Misunderstanding of terminology: acronyms being misconstrued; medical terms that may have different meanings in common use…
- Poor patient recall: patients unable to remember what operations were for and their outcomes
- Differences of opinion: about diagnosis, especially for depression; what constitutes heavy smoking; which past events are considered significant and which are listed in the patient summary
- Misunderstanding how information is managed in the NHS: the practice administrative recording system, administrative prompts, transfer of information — for example, from genito-urinary medicine clinics, out-of-hours services and emergency services

6.4.4. Furthermore, 77 from 319 patients approached declined to have or view their online record.

6.4.5. This final finding is in line with a recent survey of Walk in Centre users in an early adopter [35]. 60% of patients reported that they did not want a Health Space account, and a further 10% had not made up their minds. Of patients with low health literacy, 89% had definitely decided not to enrol for a HealthSpace account, contrasted with 56% of medium health literacy and 30% of high health literacy patients.
6.4.6. The uptake of HealthSpace accounts in practice is lower than these findings would suggest. In July this year only 597 of the 603,384 patients in receipt of an invitation to apply for an advanced HealthSpace account have applied for and activated it. (from HealthSpace Management Information, week ending: 17 July 2008).

6.4.7. There is, however, one group of patients that may be actively involved in record validation via HealthSpace. Greenhalgh et al [35] found that "For most people, the personal risk-benefit equation came out in favour of having a SCR but against having a HealthSpace account... However, a small but important minority saw great benefit in the potential of HealthSpace for keeping track of their own or a relative’s chronic illness".

6.4.8. The phrase “Expert Patient” is used to describe those patients with a reasonably stable long-term condition who are familiar enough with their condition and have the competence and engagement to play an active role in self-management. Whilst patient self management is in theory likely to encourage greater use by patients of electronic health records and therefore to enable them to correct errors, it is not yet clear what information resources in the SCR will be of value to such patients. Expert patients are potentially the group of patients with the highest health literacy and the highest motivation to access HealthSpace. These patients are the most likely group to be able to provide some validation of their SCR.

6.4.9. In general however, opt-in rates for HealthSpace are low at present. The findings [34] and [35] suggest that HealthSpace accounts are unlikely to be taken up universally. Furthermore, Pyper et al found that patients often lack the expertise to check their records. Therefore the impact of HealthSpace on data quality in the SCR is likely to be marginal.
7. Recommendations

7.1. Remit

7.1.1. CfH commissioned this evaluation in order to determine a) the effectiveness of data quality measures in place at the moment and b) the data quality measures that are desirable in order to ensure the widespread use of the SCR. The findings in this report have general applicability in all use scenarios, but we have focused on the domain of clinical use as this is where the greatest risks potentially lie.

7.2. Summary Care Record

This evaluation recommends that relevant stakeholders in the National Programme:

7.2.1. Continue to validate the quality of general practice data before it is uploaded to the SCR. The literature in this field clearly demonstrated that data quality in primary care is variable. Data quality is consistently lower for certain types of data items and, whilst most of these are not yet present on the SCR, some will be added in the foreseeable future.

7.2.2. Publish a focussed and clearly specified set of priority use cases for the SCR. This recommendation was originally made by the BCS in their 2006 report “The Way Forward for NHS Health Informatics” [6]. Having such use cases would not only serve to define the data items and use contexts for he SCR, they would also clarify the scope of each phase of the SCRs development and implementation, thus supporting both technical and change management.

7.2.3. Undertake a formal analysis of each use case, to determine the minimum utility standards that need to be met by the SCR. This analysis should include (but not be limited to):

- A definition of the set of data items required for each use case, including the data category (e.g. Most recent measured blood pressure); the data elements (e.g. systolic pressure and diastolic pressure); the data types (e.g. real valued numbers); the data units (e.g. mmHG); the allied metadata (e.g. date reading taken); level of detail (e.g. exact reading: 120 mmHG, or a categorical reading: “high”, “medium”, “low”).
- Guidelines on the limits of the utility of that data (for e.g. what are the “business rules” determining when a data item is “out of date”).

This analysis must take into account both the technical aspects of data production and use (e.g. the behaviour of recording instruments, the format of data that is to be machine-readable) and the organisational aspects of data production and use (e.g. the limits on reasonable data recording in consultation and the cognitive and behavioural limits on data access at the point of use).

7.2.4. Undertake a formal risk analysis of each use case to determine the minimum safety standards that the SCR data needs to meet. It may be that the desired production standards for given data uses are so high in practice, or so difficult to measure as to make these impossible to apply. In which case standards must be applied to the point of use.
7.2.5. Recognise that, in order to minimise the risk of clinical error, operational requirements may be applied at the point of use. Further investigations are necessary into the requirements for summary data to be safely interpreted out of the context in which it was composed, and by diverse clinical users and patients. These requirements may consist of extra validating steps at the point of use, such as mandatory triangulation with other data sources.

7.2.6. Recognise that the SCR will never be 100% complete and accurate with respect to the state of the patient. This is in part because the state of the patient will have altered since the last entry on the SCR; in part due to human fallibility and in part because data produced for one use can never be translated wholesale into another context without some losses. The potential end users we have sampled believe that the SCR will be welcomed as an additional, but not necessarily definitive, source of information. Data users must be allowed to make judgments as to the utility of the data on an as-needed basis.

7.2.7. Conduct detailed evaluations of the early experience of users of the SCR. Clinical practice, especially in unscheduled care settings, is characterised by the need for rapid decision-making. The data in SCR must be presented in a way that supports the end user in making fast, appropriate judgements about the contents of the record. Studies of the use of the SCR in practice will inform the data quality requirements and critical safety features, as well as establish good practice in the safe and appropriate use of the SCR as a contributor to informed clinical care.

7.3. Data quality validation

7.3.1. In the absence of prescriptive DQ standards that are driven by a limited, clearly defined set of use cases, it is our position that the best possible standards of data quality will be obtained through the use of ongoing audit and intervention cycles. This is a costly approach, and requires a cohort of skilled DQ facilitators that are able to engage with the practices. Nevertheless, we would argue that the audit and intervention approach is the only way to improve the clinical usability of data in the SCR. It is imperative that interventions are carried out by DQFs that have the skills and flexibility to fully engage primary care staff in order to help them change their practices.

This evaluation therefore recommends that relevant stakeholders in the National Programme:

7.3.2. Continue to recognise data quality as primarily organisational and educational processes that are ongoing and require regular stimulus. The production of good quality data relies more on the skills and professionalism of the people producing that data than it does on the systems it is recorded on. Primary care data is, by its very nature, prone to ambiguity and uncertainty. Furthermore, the data in the SCR will be used in a variety of different contexts for a variety of different reasons. Professional judgment is therefore, more appropriate than a set of prescriptive measures.

7.3.3. Extend the availability of well-trained Data Quality Facilitators to assist practices to: appraise their data quality; identify gaps and provide training in data quality improvement and monitoring and share good practice approaches. There are a number of specialist skills required in order to produce good quality data. These include technical abilities, the ability to interpret CHART results, the ability to develop good process and so on. There are also a number of professional behaviours that must be adopted in order to produce good quality data. These skills and behaviours
are unlikely to arise spontaneously from within a practice. A good relationship with a skilled data quality facilitator is required in order to help practices develop these skills.

7.3.4. Continue to support organisations and tools used to audit data completeness and accuracy in primary care, such as PRIMIS+ and CHART, and the Paperlight protocol. These tools have a useful role to play. As the SCR grows, and as the development and analysis of use cases clarifies the data quality requirements, new audits will need to be developed and existing ones refined.

7.3.5. Recognise the limited "indicator" role that established metrics of data quality play in implying but not establishing the general level of quality of practice health records, serving more as an indication of processes and training needs than an assessment of the fitness of the records for CRS upload. Metrics a) only capture data quality at one moment in time, b) are only applicable to the minority of clinical data and c) cannot capture the flexible and open-ended requirements that the SCR must address.

7.3.6. Recognise the limitations on the role of HealthSpace as a route to data quality validation. Whilst the present low rate of uptake of HealthSpace is potentially compounded by the organisational aspects of registration, research carried out by our group and others shows that the number of people competent or engaged enough to validate their own record will always be limited. A sizeable minority may eventually have a role to play in validating their own records, but the majority of patients are unlikely to engage in this way.

7.4. **Suggestions for further research**

7.4.1. In order to answer outstanding questions about the effect of data quality, it is suggested that the following research questions be addressed:

7.4.2. Regarding the production of data: what are the socio-technical factors that affect data quality as records are produced. Our research has identified examples of the use of DQFs as change agents that are able to bring about changes to work practices in order to improve data quality. A more detailed socio-technical study should take into account the changes that the SCR has wrought for primary care workers. For example, GPs must now treat their electronic records not as a personal aide memoire, but as the basis for a collaborative working relationship with other care providers and the patient. How should the GPs role change in order to address this? How has it changed already?

7.4.3. Regarding the use of data: what are the risks of poor data quality in clinical practice? This question could be answered in two ways: firstly by extrapolating from audit results and actuarial calculations of clinical risk, which would give a figure for risk that is accurate, but not necessarily relevant. The reason that this figure may not me relevant is that it would not take into account the effect of data quality on the pattern of uptake by the end users in the clinical setting. Thus we also recommend an ethnographic study, which addresses whether and how the SCR is used. Whilst it will require a higher level of resources than an actuarial approach, an ethnographic study is appropriate for the study of the risks arising from data use.

7.4.4. What are suitable evidence based quality processes and data quality metrics to establish the fitness of a practice and its records to support the defined purposes of the SCR? These metrics must be based on SCR use cases, as these will clarify how the SCR is to be used. Only then will it be possible to determine appropriate gold
standards and metrics, and how to prevent data of inappropriately low accuracy and completeness being uploaded.

7.5. **A national strategy for data quality**

7.5.1. Given the involvement of multiple stakeholders in the SCR programme, and the complexity of determining the responsibilities of these, we strongly recommend that a national strategy for data quality should be established. We commend CfH for recognising the range and complexity of a project of this magnitude. The various aspects of organisational and informatics planning must be identified and responsibility apportioned to the most appropriate stakeholders.
References


[38] Information And Statistics Division, National Health Service In Scotland, *Assessment Of Data Quality SMR01 1996/97 Scotland, Quality Assessment And Accreditation, October 1998*


## Appendix A: Glossary

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCSHIF</td>
<td>British Computer Society Health Informatics Forum</td>
</tr>
<tr>
<td>CFH</td>
<td>NHS Connecting for Health</td>
</tr>
<tr>
<td>CFHEP</td>
<td>Connecting for Health Evaluation Programme</td>
</tr>
<tr>
<td>CRS</td>
<td>Care Records Service</td>
</tr>
<tr>
<td>CSA</td>
<td>Clinical Spine Application</td>
</tr>
<tr>
<td>DMICP</td>
<td>Defence Medical Information Capacity Programme</td>
</tr>
<tr>
<td>DOH</td>
<td>Department of Health</td>
</tr>
<tr>
<td>DQF</td>
<td>Data Quality Facilitator</td>
</tr>
<tr>
<td>ECS</td>
<td>Emergency Care Summary (Scotland)</td>
</tr>
<tr>
<td>ERDIP</td>
<td>Electronic Development and Implementation Programme</td>
</tr>
<tr>
<td>EHR</td>
<td>Electronic Health Record</td>
</tr>
<tr>
<td>EMIS</td>
<td>Egton Medical Information Systems</td>
</tr>
<tr>
<td>EPR</td>
<td>Electronic Patient Record</td>
</tr>
<tr>
<td>GP</td>
<td>General practitioner</td>
</tr>
<tr>
<td>GPSoC</td>
<td>GP Systems of Choice</td>
</tr>
<tr>
<td>GPSS</td>
<td>GP software systems</td>
</tr>
<tr>
<td>HCA</td>
<td>Health care assistant</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communications Technology</td>
</tr>
<tr>
<td>LDR</td>
<td>Local Detailed Records (these include the GP held record, pharmacy record, out-of-hours record etc)</td>
</tr>
<tr>
<td>LMC</td>
<td>Local Medical Committee</td>
</tr>
<tr>
<td>LSP</td>
<td>Local Services Provider</td>
</tr>
<tr>
<td>MIU</td>
<td>Minor Injuries Unit</td>
</tr>
<tr>
<td>NCL</td>
<td>National Clinical Lead</td>
</tr>
<tr>
<td>NCRS</td>
<td>National Care Records Service</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>NLOP</td>
<td>National [Programme for IT] Local Ownership Programme</td>
</tr>
<tr>
<td>NPfIT</td>
<td>National Programme for Information Technology</td>
</tr>
<tr>
<td>OOH</td>
<td>Out of Hours</td>
</tr>
<tr>
<td>PAC</td>
<td>Public Accounts Committee</td>
</tr>
<tr>
<td>PCT</td>
<td>Primary Care Trust</td>
</tr>
<tr>
<td>PDS</td>
<td>Personal Demographic Service</td>
</tr>
<tr>
<td>PEC</td>
<td>Professional Executive Committee (of PCT)</td>
</tr>
<tr>
<td>PID</td>
<td>Project Initiation Document</td>
</tr>
<tr>
<td>QoF</td>
<td>Quality and Outcomes Framework</td>
</tr>
<tr>
<td>RCN</td>
<td>Royal College of Nursing</td>
</tr>
<tr>
<td>SCR</td>
<td>Summary Care Record</td>
</tr>
<tr>
<td>SHA</td>
<td>Strategic Health Authority</td>
</tr>
<tr>
<td>WiC</td>
<td>Walk-in Centre</td>
</tr>
</tbody>
</table>
Appendix B: Delphi Instrument

Welcome to the Summary Care Record Data Quality Delphi Panel. Many thanks for agreeing to participate in this process. For the first round of this panel, we would like you to answer, as best as you can, the questions on the following pages. We are particularly interested in the reasons for your answers, and would appreciate it if you could spend some time elaborating on the reasons for your answers.

In the second round you will receive the same questionnaire, along with an anonymous summary of the answers from your fellow participants. These answers may stimulate you to give more detail about your own views, or they may change your views.

Your answers will be anonymous, but in order to help us to analyse the data, we need to know two pieces of information:

1) What is your current role (e.g.: out of hours nurse, data quality facilitator, practice manager etc.)

2) What, if any, experience have you had with the Summary Care Record to date?

The questionnaire begins on the next page.
The Summary Care Record (SCR) contains information on current medication and allergies/reactions to drugs for all patients. In those with long term conditions such as diabetes or hypertension, it may contain additional information if the GP has chosen to add this, but you won’t know exactly what’s in there until it is opened. In the following scenarios, the SCR may be used to give more information to the doctor or nurse.

Who needs the SCR most? What order of priority would you place on making the SCR available to each of the following doctors or nurses in each of the following scenarios? For example, if you think that it is most important that the district nurse in Scenario 5 makes use of the summary care record, rank this priority 1 and say why.

Scenario 1: Junior doctor
A junior doctor is in working the Accident and Emergency Department and is called urgently to see a woman who has been brought in unconscious. Her teenage son says that she is 42, and that his dad left home last week after a row. The woman is deeply unconscious, very pale, and smells strongly of alcohol. The boy says that the GP put his mother on some tablets last week but he couldn’t find them at home and does not know what they are.

Scenario 2: Nurse in walk in centre
A nurse is working in a walk-in centre, and sees a 68 year old woman who appears to have mild learning difficulties. She has scraped her leg against a dirty fence post while gardening, and is bleeding quite heavily from her shin. The patient says that she once had an episode where her lips and throat swelled up. She thinks this might have happened after an injection in hospital, but she can’t remember what the injection was. In any case, she hasn’t had any injections for over 20 years. The nurse must decide whether the patient should have a tetanus shot.

Scenario 3: Triage nurse
A triage nurse is working at an out of hours centre. It is Bank Holiday Monday and there is a long queue of patients waiting to see the only doctor on duty. Mr Fernando, a 56 year old gentleman, has come because of a stomach ache. He says he got some tablets from his GP last week and they haven’t helped. In fact he thinks they have made him worse. Mr Fernando’s wife says she doesn’t think he’s very well at all: he’s been weak and tired recently, but she seems a bit of a ‘pushy relative’ so maybe she’s just trying to jump the queue. However, Mr Fernando looks in quite a bit of pain. The triage nurse needs to determine if he should be flagged as ‘urgent’.

Scenario 4: General Practitioner (Out of Hours)
Mr Fernando in the above scenario has been referred to an out of hours GP who has examined him and found that he has a tender abdomen. His pulse is slightly rapid (90) and his blood pressure is normal. The GP must decide whether to send him into hospital for further investigation, but apparently his mother is visiting for a few days so he’d rather just have some new medicine and go home.

Scenario 5: District Nurse
A district nurse doing a routine weekly visit to Mrs Begum, a 73 year old Bangladeshi lady with diabetes and leg ulcers. Her daughter-in-law says Mrs Begum has uncharacteristically taken to her bed and has been there for three days and isn’t steady enough on her feet to walk to the toilet. Mrs Begum is lying in bed, apparently asleep. When you try to rouse her, she groans and opens her eyes briefly. You’re not really sure if she’s got a decreased level of consciousness or if she’s just sleepy (or being difficult). Apparently she had a full diabetes check at the GP surgery three weeks ago. Mrs Begum’s tablets, of which there are several types, are laid out in a dose-it box. A finger prick for blood glucose is slightly above normal. The district nurse must decide whether or not Mrs Begum should be taken to the (very busy) hospital.

Scenario 6: NHS Direct
A nurse at NHS Direct receives a phone call from the partner of Ms Brown aged 27, who is recovering from a termination of pregnancy operation 4 days ago. She is apparently feeling dizzy and sick. Ms Brown suffers from mild epilepsy, for which she takes sodium valproate, a drug that may interfere with the contraceptive pill. She visited the GP yesterday, who gave her some painkillers and started her on a contraceptive pill for the first time. She took the first pill this morning. The NHS Direct nurse must give advice.
Imagine that you have information as to the completeness and accuracy of the data about current medications and allergies on the SCR. If you knew that:

- 90% of the SCRs are complete - that is, all data that should be there is on there and
- 98% of the SCRs are accurate - that is, all data that is on there is correct

Who needs the SCR most?

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Priority (1-6)</th>
<th>Reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Junior doctor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2: Nurse in walk in centre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: Triage nurse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: GP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5: District nurse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6: NHS Direct</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Imagine now that you knew that:

- 70% of the SCRs are complete - that is, all data that should be there is on there and
- 90% of the SCRs are accurate - that is, all data that is on there is correct

How might your priorities change? If there has been a change in priority, please say why.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Priority (1-6)</th>
<th>Reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Junior doctor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2: Nurse in walk in centre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: Triage nurse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: GP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5: District nurse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6: NHS Direct</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Explaining differences in data quality
Anystreet Practice has excellent data quality: audits and manual inspections have shown that the data entered on the SCR are complete (all allergies and drugs for each patient are entered onto the record) and accurate (all allergies and drugs on the record are actually suffered/taken by the patient). This has taken a lot of work by the practice staff. Otherstreet Surgery however is struggling to achieve a minimum data quality standard. Many entries that would be expected on the electronic record are missing, and there are known to be patient records in which drugs or allergies have been entered on the wrong patient's record.

Anystreet and Otherstreet operate in the same area and cater for the same type of patient mix. Nevertheless, their performances in relation to data quality are extremely different. Why has Anystreet Practice got such good data quality? Why might Otherstreet be finding this so much more challenging?

- Who is likely to be leading on data quality in each of these two practices, and how are they going about this?

- What are the main motivations of the staff in these two practices?

- What other priorities or issues might these practices have, that they need to balance with data quality?

- Any other differences?