

---

# **GMC Multi-Source Feedback Questionnaires**

## **Interpreting and handling multisource feedback results: Guidance for appraisers**

Professor John Campbell (Academic Lead)

Dr Christine Wright (Research Fellow)

Primary Care Research Group, Peninsula College of Medicine & Dentistry,  
Smeall Building, St Luke's Campus, Exeter, United Kingdom, EX1 2LU.

**Guidance prepared: 14 October 2011**



---

## Research Team and Acknowledgments

This guidance has been derived from the findings of a programme of research led the Primary Care Research Group, Peninsula College of Medicine and Dentistry, Exeter. Core members of the research team included:

Professor John Campbell (Academic Lead)<sup>1</sup>.

Ms Jacqueline Hill (Research Fellow)<sup>1</sup>.

Dr Suzanne Richards (Senior Lecturer)<sup>1</sup>.

Mr Martin Roberts (Statistician, Research Fellow)<sup>1</sup>

Dr Christine Wright (Research Fellow)<sup>1</sup>.

- <sup>1</sup>. Primary Care Research Group, Peninsula College of Medicine & Dentistry, Smeall Building, St Luke's Campus, Exeter, EX1 2LU.

We acknowledge the extensive support provided throughout the research programme of the GMC-commissioned partner survey organisation, CFEP-UK: Professor Michael Greco was instrumental in the development of the work from its earliest stages. Mr Matthew Taylor and Ms Louise Coleman managed the recruitment of doctors and coordinated the multisource feedback data collection. Mrs Helen Powell helped to develop the feedback reports.

Dr Julian Archer (Peninsula College of Medicine & Dentistry, Plymouth) has also acted as consultant adviser to the authors and we thank him for his valuable advice and feedback in the development of these guidelines.

---

## Contents

1	Introduction .....	5
1.1	Background .....	5
1.2	Purpose of the guidance.....	6
2	Description of the GMC questionnaires .....	8
2.1	GMC Patient Questionnaire (PQ) .....	8
2.2	GMC Colleague Questionnaire (CQ) .....	9
2.3	GMC Self-assessment Questionnaire (SQ) .....	10
3	The GMC Feedback Report.....	11
3.1	Patient and colleague feedback sections.....	11
3.2	Self-assessment section.....	15
3.3	Explanatory materials and supporting documents .....	16
4	Findings of recent studies .....	17
4.1	Skewed nature of MSF data .....	17
4.2	Reliability of MSF results .....	18
4.3	Item completion rates .....	18
4.4	Volunteer nature of benchmark data .....	19
4.5	Effects of patient characteristics on item scores.....	19
4.6	Effects of colleague characteristics on item scores .....	20
4.7	Effects of doctor characteristics.....	21
5	Reviewing and interpreting the MSF report.....	23
5.1	Independent reflection on MSF results .....	23
5.2	The introductory text.....	23
5.3	The patient feedback section.....	24
5.4	The colleague feedback section .....	33
5.5	Self-assessment section.....	43

6	Determining what further action is needed .....	45
6.1	Making judgements about the need for further action.....	45
6.2	Managing discussions about further action .....	45
6.3	Deciding on further action: what about ‘low performers?’ .....	46
6.4	Serious concerns.....	48
7	References .....	50
	Appendix 1: Content of the GMC Patient Questionnaire.....	53
	Appendix 2: Content of the GMC Colleague Questionnaire.....	55

**List of figures**

Figure 1:	Patient ratings on the GMC Patient Questionnaire (PQ) .....	28
Figure 2:	Colleague ratings on the GMC Colleague Questionnaire (CQ).....	37

---

# 1 Introduction

## 1.1 Background

All doctors who wish to practise medicine in the United Kingdom need to be registered with and licensed by the General Medical Council (GMC). To retain their licence with the GMC, doctors will also be required to revalidate on a regular basis to demonstrate that they are still 'fit to practise' medicine.

The GMC's proposals for the revalidation process have now been published.<sup>1-3</sup> As part of that process, doctors will need to collate 'supporting information'<sup>4</sup> to demonstrate they meet relevant standards of professionalism. Such information will feed into the appraisal process,<sup>3,4</sup> and ultimately lead to a recommendation by a Responsible Officer to the GMC about the doctor's suitability for revalidation. The collection of feedback from colleagues and patients, using structured questionnaires is likely to be an integral component of the revalidation process for doctors.<sup>4,5</sup>

Early reviews<sup>6,7</sup> of the tools available for this purpose (conducted in 2004 and 2006) suggested that few had undergone sufficiently rigorous testing of reliability and validity. In the absence of suitably robust instruments to assess the professional performance of doctors, the GMC developed its own patient and colleague questionnaires, with content based on the principles of 'Good Medical Practice'.<sup>8</sup> A series of early pilot studies explored their utility, reliability and validity.<sup>9-11</sup> These studies suggested the questionnaires were acceptable to respondents, offered preliminary reassurance regarding their validity, reliability and feasibility, and indicated they could identify a range of performance in respect of professionalism in a volunteer sample of doctors.

In 2010, a report commissioned by the Royal College of General Practitioners<sup>12</sup> reviewed nine colleague feedback instruments and ten patient feedback instruments. That report concluded that considerable work had been undertaken to develop questionnaires that might be capable of assessing the professional performance of doctors for the purposes of revalidation. In particular, the review noted that the GMC Patient Questionnaire (PQ) and the GMC Colleague Questionnaire (CQ) mapped well onto the competencies required for 'Good Medical Practice'<sup>8</sup> and that comprehensive psychometric testing, in line with the GMC's Framework Attributes,<sup>3</sup> had been undertaken.

The GMC questionnaires have recently undergone a further phase of development and evaluation using a large sample of UK doctors working across a wide range of NHS and non-NHS clinical settings. This phase addressed some of the gaps in previous development work, including an exploration of the test-retest reliability and convergent validity of the PQ and CQ, as well as the effect of patient, colleague and doctor characteristics on ratings obtained via these questionnaires. Confidential reports were prepared between November 2010 and March 2011 for consideration by the GMC and their technical advisers. A series of academic papers have also been prepared for publication in scientific journals.<sup>13-17</sup>

## **1.2 Purpose of the guidance**

This document provides a brief description of the GMC multisource feedback (MSF) questionnaires (Chapter 2), as well as the personalised report issued to doctors at the end of the survey process (Chapter 3). A review of the findings of recent pilot work is provided in so far as these relate to the interpretation of the

MSF results (Chapter 4). Guidance is provided for appraisers on how they might approach the review and interpretation of a doctor's MSF report (Chapter 5), and decide what further action (if any) is required based on the feedback (Chapter 6).

Doctors and appraisers may also wish to refer to the GMC's guidance<sup>4</sup> on how each element of the supporting information collected by doctors for the purposes of revalidation might be used or discussed in appraisal. This is available at:

[http://www.gmc-uk.org/doctors/revalidation/supporting\\_information.asp](http://www.gmc-uk.org/doctors/revalidation/supporting_information.asp)

---

## **2 Description of the GMC questionnaires**

For most doctors, the process for collecting feedback on their practice will include a patient survey and a colleague survey. Doctors whose current role does not include direct consultations with patients will complete only the colleague survey. The GMC has specified the principles and criteria that all patient and colleague questionnaires must meet for the purposes of revalidation.<sup>18</sup>

The GMC has developed questionnaires which might be used to collect such feedback – namely the GMC Patient Questionnaire (PQ) and the GMC Colleague Questionnaire (CQ). A self-assessment questionnaire (SQ) is also available. The questionnaires include items relating to the GMC's core guidance on the principles and values to which it requires registered doctors to adhere.<sup>8</sup> The content of the PQ and the CQ is summarised in Appendices 1 and 2.

### **2.1 GMC Patient Questionnaire (PQ)**

The PQ comprises 9 core items which assess the doctor's consultation skills and aspects of their probity. Other items collect information about the context in which the questionnaire has been completed, and about the patient. If the patient is a child aged 12 years or less, lacks mental capacity, or is too ill or disabled to complete the questionnaire, a carer (or 'proxy') can complete it on the patient's behalf.

The questionnaire is designed to be administered to 45 consecutive patients (or carers) as a post-consultation or 'exit' survey. Recent evidence<sup>14</sup> suggests that, to achieve reliable results, a minimum of 34 PQs need to be returned. This

figure will be further investigated as evidence accumulates relating to the use of the PQ in practice.

In clinic settings, the survey pack can be distributed by reception staff or other clinic staff (e.g. nurses). Patients are encouraged wherever possible to complete their questionnaire in the waiting area, immediately after their appointment with the doctor, and return it in a sealed envelope to a ballot box in the clinic.

In other settings (e.g. anaesthetics), the doctor may need to approach patients and distribute the survey packs themselves at the end of the consultation. If a ballot box is not feasible, patients may take the questionnaire home and post their completed questionnaire to the survey organisation in a reply-paid envelope.

## **2.2 GMC Colleague Questionnaire (CQ)**

The CQ comprises 19 core items which assess the doctor's clinical, communication, organisational and teaching skills as well as aspects of their probity and health. Other items collect information about the colleague respondent and their familiarity with the doctor's practice.

At the start of the MSF process, the doctor is asked to nominate 20 colleagues (10 medical; 10 non-medical) who are able to provide feedback on their professional performance. Recent evidence<sup>14</sup> suggests that, to achieve reliable results, a minimum of 15 CQs need to be returned. This figure will be further investigated as evidence accumulates relating to the use of the CQ in practice.

Data collection for the colleague survey is managed by an independent survey organisation. Doctors provide the survey organisation with a list of their nominated colleagues, and their e-mail or postal addresses.

Each colleague on the list is approached by the survey organisation and invited to complete an online CQ, using a unique log-in and password. A paper version of the CQ is available on request. Two reminders are issued to non-responding colleagues.

### **2.3 GMC Self-assessment Questionnaire (SQ)**

Previous research suggests that self-assessment is a cornerstone of self-directed professional development<sup>19-21</sup> and that disagreement with negative feedback can affect the likelihood that doctors will act on such feedback.<sup>22-24</sup>

Furthermore, a lack of insight into one's performance appears to be more prevalent amongst certain groups of individuals, including those who are at the lower end of the performance spectrum.<sup>25-26</sup>

The GMC has developed a self-assessment questionnaire (SQ) which comprises 26 core items, and maps on to the content of the PQ (7 items) and the CQ (19 items). Ten other items collect background information about the doctor and the context in which they practice.

Each doctor who undertakes the MSF process is invited by the survey organisation to complete an online SQ, using a unique log-in and password. A paper version of the SQ is available on request. Two reminders are issued to non-responding doctors.

---

## **3 The GMC Feedback Report**

On completion of the patient and colleague survey processes, doctors receive a personalised report which summarises their results. Currently, the report is sent directly to the doctor from the survey organisation.

At the beginning of the survey process, doctors are encouraged to nominate a 'supporting medical colleague' with whom they could informally discuss their report shortly after its receipt. The nominated supporting medical colleague is notified by the survey organisation when the doctor's report is issued, but they are not sent a copy of the report. It is entirely up to the doctor to decide whether they wish to make use of this support mechanism when their report is received. Doctors are also encouraged to discuss the report with their appraiser.

The template for the feedback report has undergone two waves of evaluation and revision, during which the views of doctors were sought and used to develop the report template.<sup>27</sup> The GMC feedback report is divided into three key data sections – patient feedback, colleague feedback and self-assessment.

### **3.1 Patient and colleague feedback sections**

The patient and colleague sections of the report follow a similar format, which includes:

#### **3.1.1 *Information about the survey samples***

The report provides background information about the samples of patients and colleagues who have taken part in the survey – e.g. with regard to their gender, age group and (colleagues only) professional group. To maintain anonymity of

participants, if there were less than three respondents in a particular sub-group of age, gender or professional group, the sub-group is not included in the table.

Your patient feedback is based on responses from 22 patients with the following characteristics:

Table 1.1:

Gender

Female	8	36%
Male	14	64%

Age

Under 15	8	36%
15-20	14	64%

Number and percentage of responses by question (percentage of responses may not add up to 100% due to rounding). To maintain anonymity of participants, if there are less than three responses in any category in gender or age, that category is not reported.

### 3.1.2 Frequency and distribution tables

Information about the frequency and distribution of patient responses across the 9 core items of the PQ, and the frequency and distribution of colleague responses across the 19 core items of the CQ is presented.

## GMC Patient Feedback Report

Number of patients providing feedback: 22

### Evaluation question ratings and scores

Table 1.2: Distribution and frequency of ratings (Q4 - Q8)

	Poor	Less than satisfactory	Satisfactory	Good	Very good	Does not apply	Blank (missing)	Spoilt
Q4a Being polite	0	0	1	0	21	0	0	0
Q4b Making you feel at ease	0	0	1	1	18	0	2	0
Q4c Listening to you	0	1	1	3	17	0	0	0
Q4d Assessing your medical condition	0	2	4	9	7	0	0	0
Q4e Explaining your condition and treatment	0	2	4	9	7	0	0	0
Q4f Involving you in decisions	0	2	4	9	7	0	0	0
Q4g Providing or arranging treatment for you	0	2	4	9	7	0	0	0

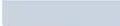
The frequency of 'Does not apply' responses and missing or spoilt data for each item is also included.

### 3.1.3 Benchmark data

For each of the 7 PQ items and each of the 18 CQ items that are rated on 5-point scales, a mean percentage score is calculated, using a process that is described in the supporting documents at the end of the report.

GMC Patient Feedback Report		Number of patients providing feedback: 22				
Evaluation question ratings and scores						
Table 1.3: Mean percentage scores and benchmarks (Q4,Q5)						
	Your mean score (%)	Benchmark data (%) *				
		Min	Lower quartile	Median	Upper quartile	Max
Q4a Being polite	98	70	96	98	99	100
Q4b Making you feel at ease	-	-	-	-	-	-
Q4c Listening to you	91	61	94	97	98	100
Q4d Assessing your medical condition	74	68	93	96	98	100
Q4e Explaining your condition and treatment	74	65	93	95	98	100
Q4f Involving you in decisions	74	67	92	95	97	100
Q4g Providing or arranging treatment for you	74	68	93	96	98	100
Q5a Confidentiality of information	74	59	90	93	95	100
Q5b Doctor is honest and trustworthy	74	65	91	94	96	100

	Your mean score for this question falls in the highest 25% of all means
	Your mean score for this question falls in the middle 50% of all means
	Your mean score for this question falls in the lowest 25% of all means

- insufficient number of responses to generate score  
See score explanation for percentage score calculation and quartile information

The doctor's mean percentage score on each of the core items is presented alongside corresponding item benchmark data – i.e. minimum score, lower quartile, median, upper quartile and maximum percentage score values achieved on that item by doctors who have so far completed the MSF process. A colour coding system indicates where the doctor's mean percentage score falls in relation to those of other doctors (e.g. in the lowest 25% doctors; middle 50% doctors, or highest 25% doctors).

Notes about the benchmark data appear underneath each of the tables and it is strongly recommended that doctors and their appraisers read this information before attempting to interpret the data in the relevant table.

**\* Important notes about the benchmark data**

The above benchmarks are based on all data that is available for the doctors who participated in our pilot work up to 2 July 2009. To calculate the benchmarks for each questionnaire item, we have included data only for doctors who had at least six valid responses returned for that particular item. We will update these benchmarks at regular intervals as the number of doctors undertaking these surveys increases.

Please bear in mind the following points when interpreting the above table:

**Purpose:** These benchmarks are provided to give you a sense of how you are performing in relation to other doctors who have completed the GMC surveys. They are not intended to imply any 'minimum standard' that doctors are expected to achieve for the purposes of revalidation.

**Sample size:** Of the doctors who have agreed to take part in the pilot work so far, 935 have completed a patient survey and 942 have completed a colleague survey. The benchmark data contained in this report is derived only from the doctors who returned sufficient numbers of patient or colleague responses. To date, 560 doctors have returned sufficient patient feedback and 620 doctors have returned sufficient colleague feedback.

**Voluntary participation:** It is not mandatory for any doctor to undertake these surveys. Therefore the benchmarks are based on a volunteer sample of doctors. As such, the benchmarks may be higher than might be expected if all doctors had contributed data.

**Range of practice:** The doctors who have contributed to the benchmark data work or provide care in a variety of settings and specialties. The above benchmarks relate to the whole sample of doctors who have contributed data, irrespective of their setting or specialty.

Currently, the GMC feedback report provides two types of benchmark data: 'generic' benchmarks and 'setting-specific' benchmarks. Generic benchmarks are derived from all doctors, *regardless* of their clinical setting, while setting-specific benchmarks are derived from doctors who practise in the *same* clinical setting as the doctor (i.e. primary care or secondary care).

The report also provides web links to the developing 'specialty-specific' benchmarks (currently available for medicine, surgery, psychiatry and general practice only). In future, as the number of doctors completing the MSF process increases, it should be possible to provide robust benchmark data that is specific to a wider range of clinical specialties.

### 3.1.4 Free text comments

The report includes any free text comments made on the questionnaires by the doctor's patients and colleagues. All comments are screened before reporting, to remove any text which might identify the respondent. Whilst previous research<sup>28</sup> suggests that free text comments on the CQ do not contain any information that is not already captured by the rating scales of the questionnaire, some doctors have found the comments useful in terms of understanding the numerical ratings in earlier sections, and helping to identify specific ways in which their performance might be improved.<sup>27</sup>

## 3.2 Self-assessment section

In this section, the doctor's self-assessment ratings are presented alongside the mean score (range 1-5) obtained from patients and colleagues on each PQ and CQ core item. This section allows the doctor to compare their own views of their professional performance with the views of their patients and colleagues.

### GMC Self Assessment Report

Number of patients providing feedback: 22  
Number of colleagues providing feedback: 8

Comparison of self assessed scores with patient and colleague scores

Table 3.1:

Scores provided on a 1 - 5 scale where 1=Poor, 2=Less than satisfactory, 3=Satisfactory, 4=Good and 5=Very good

Patient questions	Your assessment	Patient assessment *
Q3a Being polite to patients	5	4.9
Q3b Making patients feel at ease in your presence	5	-
Q3c Listening to patients	5	4.6
Q3d Assessing patients' medical conditions	5	4.0
Q3e Explaining patients' conditions and treatment	5	4.0
Q3f Involving patients in decisions about treatment	5	4.0
Q3g Providing or arranging treatment for patients	5	4.0

### **3.3 Explanatory materials and supporting documents**

For transparency and to facilitate the interpretation of the doctor's survey data, the report includes sections of text which explain in detail how the mean percentage scores are calculated, and how the benchmark data has been derived, as well as an explanation of the quartile bandings. A copy of the PQ and CQ also appear as appendices to the report.

Draft

---

## 4 Findings of recent studies

This chapter summarises the findings of recent studies<sup>9,13,14</sup> insofar as these relate to understanding and interpreting patient and colleague feedback obtained via the GMC questionnaires. A number of important points should be borne in mind by doctors and appraisers when they are reviewing and interpreting MSF results.

### 4.1 Skewed nature of MSF data

In pilot work conducted in 2008-2010 with a sample of 1065 non-training grade doctors,<sup>14</sup> patient and colleague responses on the core items of the GMC questionnaires tended to be overwhelmingly positive. This pattern of response was also found in previous pilot work.<sup>9</sup>

In the 2008-2010 pilot work, the majority of the 30,333 patients who responded rated their doctor's performance as 'Very good' or 'Good' on the core items of the PQ (84% to 98% patients, varying across items). Only 1% or fewer patients rated their doctor's performance as 'Less than satisfactory' or 'Poor'. Similarly, on the summative items, the majority of patients (98%) indicated that they were confident in the doctor's ability to provide care and would be happy to see the doctor again.

There was also evidence of a bias towards positive assessments from the 17,012 colleagues who responded. For example, the majority of colleagues rated the doctor's performance as 'Very good' or 'Good' on the core items of the CQ (68% to 98% varying across items); only 1% or fewer colleagues rated the doctor's performance as 'Less than satisfactory' or 'Poor'. On the summative

item, the majority of colleagues (97%) agreed that the doctor was fit to practise medicine.

## **4.2 Reliability of MSF results**

The recent pilot work<sup>14</sup> suggests that at least 34 completed PQs and 15 completed CQs are needed for reliable results<sup>a</sup>. Below these levels, the reliability of the results cannot be guaranteed, even for formative feedback.

## **4.3 Item completion rates**

In previous pilot work, item completion rates on both questionnaires were good. Missing or spoilt data on the returned questionnaires (1% to 3% across the core PQ items; <1% on the core CQ items) was minimal. The proportion of 'Does not apply' (PQ) and 'Don't know' (CQ) responses varied across the core items of the questionnaires.

In the colleague surveys, use of the 'Don't know' response category varied across the different professional groups but appeared logical. This suggests that, in general, respondents do not simply 'guess' at a rating where they lack the necessary expertise to make an assessment of the doctor's performance or do not have the opportunity to observe that aspect of the doctor's performance. For example, administrative staff and non-clinical managers were more likely than medical colleagues or other health care professionals to use the 'Don't know' response option on core items that related to clinical practice (e.g. clinical knowledge, diagnosis, clinical decision making, treatment, and prescribing).

---

<sup>a</sup> Reliability was assessed using a decision (D) study. In view of the pragmatic nature of the sampling, undertaken with untrained patient and colleague assessors, a threshold of  $G=0.70$  was adopted in line with expert opinion.<sup>29</sup>

#### **4.4 Volunteer nature of benchmark data**

Currently the GMC benchmark data is derived from survey data collected by volunteer samples of doctors who are likely to represent the higher end of the range of doctor performance, rather than the full range of doctor performance.

As a result, the differences between the lower quartile (LQ) and upper quartile (UQ) cut-off points in the benchmark data tend to be small for all the core items of both questionnaires. Across the PQ core items, the range of the differences between the LQ and the UQ values ranges from 3% to 5%. Across the CQ core items, the range of the differences between the LQ and the UQ values range from 4% to 11%.

This pattern may of course change as the MSF process becomes a compulsory part of the revalidation cycle and more doctors contribute data from which the benchmarks can be updated. However, at present, an item mean percentage score that falls in the lowest 25% doctors (LQ) need not necessarily mean the doctor's performance is poor or less than satisfactory. Even a small number of 'satisfactory' ratings can potentially place the doctor in the lower quartile band.

#### **4.5 Effects of patient characteristics on item scores**

The recent pilot work<sup>13,14</sup> suggests that a number of patient characteristics can influence the ratings that doctors receive on the individual items of the PQ. If the doctor has a high proportion of patients from these groups in their survey sample, it could affect the ratings they have obtained.

Based on the available evidence, some of these patient characteristics can be classified as being of "definite importance" (i.e. appear to affect ratings even

when taking account of the characteristics of the doctor who is being assessed), whilst others may be classified as being of “possible importance”.

#### **4.5.1            *Characteristics of ‘definite importance’***

##### ***(a) Perceived importance of the consultation***

Patients who perceive their visit to the doctor to be ‘very important’ tend to give higher ratings than patients who do see the visit as less important.

##### ***(b) Established doctor-patient relationship***

Patients who reported seeing their ‘usual doctor’ tend to give higher ratings than those who were not seeing their usual doctor.

##### ***(c) Patient ethnic origin***

Patients from White ethnic backgrounds tend to give more favourable ratings (on some items) than those from other ethnic groups.

#### **4.5.2            *Characteristics of ‘possible importance’***

##### ***(a) Survey administration method***

Patients who return their PQ via post, rather than via a ballot box in a clinic, tend to provide less favourable ratings (on some items).

##### ***(b) Patient age***

Older patients (aged 40 and above) tend to give more favourable ratings (on some items) than younger patients.

## **4.6    *Effects of colleague characteristics on item scores***

The recent pilot work<sup>13,14</sup> also suggests that characteristics of the colleague sample can influence the ratings that doctors receive on the individual items of the CQ. If a doctor has a high proportion of colleagues from these groups in their sample, it could affect the ratings they have obtained.

Based on the available evidence, some of these colleague characteristics can be classified as being of “definite importance” (i.e. appear to affect ratings even when taking account of the characteristics of the doctor who is being assessed), whilst others may be classified as being of “possible importance”.

#### **4.6.1            *Characteristics of ‘definite importance’***

##### ***(a) Frequency of contact***

Colleagues who have contact with the doctor on most working days tend to give more favourable ratings than those who have less than monthly contact with the doctor (on some items).

#### **4.6.2            *Characteristics of ‘possible importance’***

##### ***(a) Professional role***

Colleagues who have managerial or administrative roles, and health professionals in non-medical roles tend to give more favourable ratings than medical colleagues (on some items).

##### ***(b) Survey administration method***

Colleagues who completed their CQ online tended to have more favourable views than those who completed a paper version of the CQ – on one item only (the doctor’s respect for confidentiality).

### **4.7      *Effects of doctor characteristics***

Recent analysis<sup>13</sup> suggests that characteristics of the doctor may also influence the ‘summary scores’ they achieve on the PQ and the CQ. At the moment, doctors do not receive a note of their ‘summary scores’; they only receive mean percentage scores for individual items. This is an area of the reports which will be kept under review.

#### **4.7.1 Primary medical degree**

Doctors who obtained their primary medical degree from any non-European country, tend to receive less favourable feedback from patients than those qualifying in Europe. Doctors who obtained their primary medical degree outside of the UK or South Asia, tend to receive less favourable feedback from colleagues than doctors qualifying in those two regions.

#### **4.7.2 Clinical specialty**

Doctors who practise as a psychiatrist tend to receive less favourable feedback from patients than doctors working in other clinical specialties. Doctors practising as a general practitioner or a psychiatrist tended to receive less favourable feedback from colleagues than doctors working in other clinical specialties.

#### **4.7.3 Contractual role (grade)**

Doctors who are employed in a contractual role (grade) *other* than a general practitioner or a consultant – for example, associate specialists – tend to receive less favourable feedback from their colleagues.

#### **4.7.4 Locum status**

Doctors who are working in a locum capacity tend to receive less favourable feedback from colleagues than doctors in permanent positions.

---

## 5 Reviewing and interpreting the MSF report

This chapter provides suggestions to guide the review and interpretation of MSF reports in the context of appraisal. The approach outlined should help doctors and appraisers to gain a better understanding of the MSF results. It is important to remember that MSF results are intended to be **formative** in nature, rather than summative. For the purposes of revalidation, and within the formal appraisal process, the MSF results should be considered alongside the full range of other evidence that the doctor collects during each five-year revalidation cycle.

### 5.1 Independent reflection on MSF results

The GMC feedback report is sent directly to the doctor and they are encouraged to spend time reviewing and reflecting upon their results. Therefore, appraisers should expect that doctors have carefully considered their MSF results prior to their appraisal meeting.<sup>4</sup> Some doctors may also have discussed their results with another medical colleague.

Recent research<sup>27</sup> suggests that, during individual review and reflection, doctors may only scan quickly through the report and hone in on individual ratings, scores or free text comments that stand out, particularly those that appear more 'negative'. This may mean that too much emphasis can be placed on such feedback in the doctor's personal interpretation of the report.

### 5.2 The introductory text

Before viewing the data tables and free text comments, it is advisable to read the explanatory material. Many doctors who took part in a recent qualitative

study<sup>27</sup> reported they had skipped over the introductory text. Some doctors reported being confused by or alarmed at their MSF results, particularly with regard to the benchmark bandings. However, once the doctors returned to the introductory text and read this more carefully, they reported having a better understanding of their results.

## 5.3 The patient feedback section

It is recommended that the tables in the patient feedback section are reviewed in the order in which they are presented.

### 5.3.1 *Information about the patient sample*

As an initial step, the doctor and appraiser may wish to reflect on the number of questionnaires returned, the way in which the survey was carried out, and the characteristics of the sample. The following aspects might be discussed:

#### **(a) *The number of patients surveyed***

- ✘ How many questionnaires were handed out to patients?
- ✘ Have sufficient patient questionnaires been returned?

- ① *Doctors are supplied with 45 patient surveys and asked to distribute them all.*
- ① *34 or more completed questionnaires are required to ensure that the results are sufficiently reliable for use in formative feedback.*
- ① *Some doctors may find it difficult to collect sufficient patient feedback – e.g. those who work intensively over a long period of time with a relatively small caseload, or who work mainly with patients who are critically ill or lack mental capacity.*

#### **(b) *How the survey was conducted***

- ✘ Were questionnaires handed out to consecutive patients who consulted the doctor, or only to patients who attended particular clinics or wards?

① The GMC's guidance (2011)<sup>4</sup> for doctors states:

*“The exercise should reflect the whole scope of your practice. The range of patients providing feedback should reflect the range of patients that you see.”*

- ✘ Were the doctor's patients likely to consider their consultation was 'very important' or believe they were seeing their 'usual doctor'?
- ✘ Did patients return their completed questionnaires to a ballot box in the clinic/ward, or did they post them back to the survey organisation?

- ① *Patients who believe their consultation is 'very important' or report seeing their 'usual doctor' tend to give more favourable feedback.*
- ① *Patients who return their questionnaire via post tend to give less favourable feedback than those who return the questionnaire to a ballot box in the clinic.*

**(c) The characteristics of the patient sample**

- ✘ Is the sample representative of the patients who usually consult the doctor in terms of their age and gender?
- ✘ Are the views of particular sub-groups of patients overrepresented in the survey?

- ① *Older patients (i.e. aged 40 and over) tend to give more favourable feedback than younger patients.*
- ① *Patients from 'White' ethnic backgrounds tend to give more favourable feedback than those from other ethnic groups.*
- ① *Ratings of patients do not differ significantly from ratings of carers.*
- ① *Ratings do not appear to be influenced by the gender of the patient.*

The doctor and appraiser may wish to consider whether any of the above factors may have affected the results obtained in the patient survey.

### 5.3.2 *Frequency and distribution table*

To obtain a balanced view of the feedback obtained, the doctor and the appraiser should review and reflect on the table that presents the frequency and distribution of patient ratings. The following aspects might be considered:

#### **(a) The proportion of 'valid' responses**

- ✘ What level of missing/spoilt data does the doctor have for each item?
- ✘ What proportions of patients used the 'Does not apply' response option for each item?

In recent pilot work:<sup>14</sup>

- ① *Missing and spoilt data was minimal, ranging from 1% to 3% across the 11 core PQ items.*
- ① *Use of the 'Does not apply' option was minimal (1% to 2% patients) for most items but slightly higher (4% to 11% patients) for the items which refer to aspects of treatment (4e to 4g).*

- ✘ On each item, how many 'valid' responses are there?

Valid responses utilise the following options on the rating scale and exclude 'Does not apply' responses, and missing or spoilt data:

- ① *'Poor' to 'Very good' (Items 4a to 4g).*
- ① *'Strongly disagree' to 'Strongly agree' (Items 5a and 5b).*
- ① *'Yes' or 'No' (Items 6 and 7)*

#### **(b) The distribution of 'valid' responses**

For each item on the PQ, the doctor and the appraiser might wish to identify:

- ✘ The range (or spread) of responses across the scale.
- ✘ The response option(s) that patients have used most commonly to rate the doctor's performance.
- ✘ Whether the ratings are mainly positive, neutral or negative.

### **(c) Identifying possible areas of relative strength and weakness**

Having reviewed the frequency and distribution of scores, the doctor and appraiser may wish to discuss:

- ✘ Does the distribution of patient ratings vary across the different items on the questionnaire?
- ✘ Are there any obvious areas of strength or weakness in the doctor's performance based on the patient ratings?

The GMC's guidance for doctors<sup>4</sup> recommends that:

- ① *"The discussion at appraisal should highlight areas of good performance and help you to identify any areas that might require further development".*

It may be helpful to compare the distribution of the doctor's patient ratings with the distribution of patient ratings in recent pilot work<sup>14</sup> (see Figure 1).

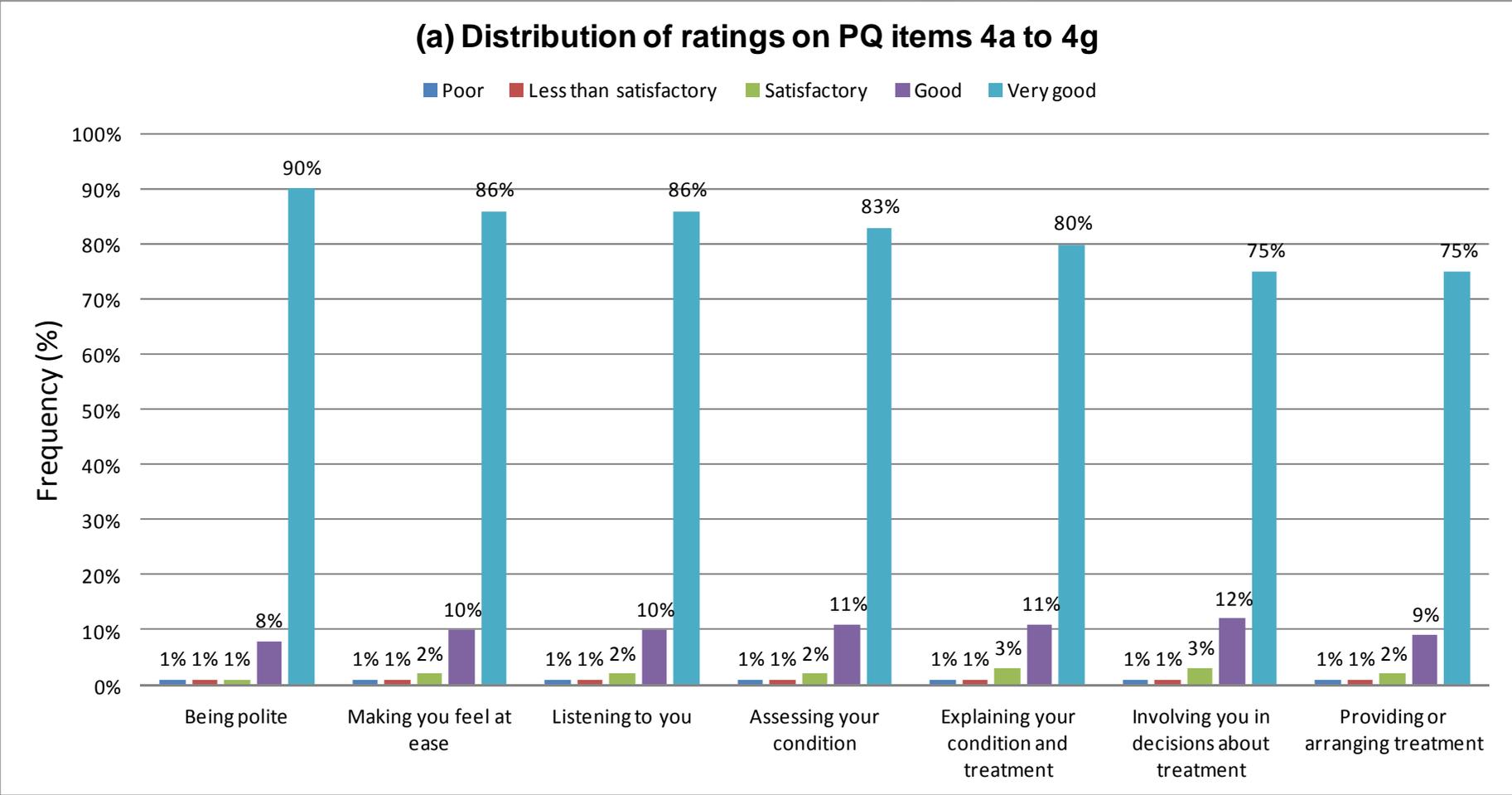
For example, on Item 4a of the PQ ('Being polite'):

- ① *Most patients in the pilot work selected the 'Very good' (90%) or 'Good' (8%) response options;*
- ① *Smaller proportions of patients selected the 'Satisfactory' (1%), 'Less than satisfactory' (1%), or 'Poor' (1%) response options.*

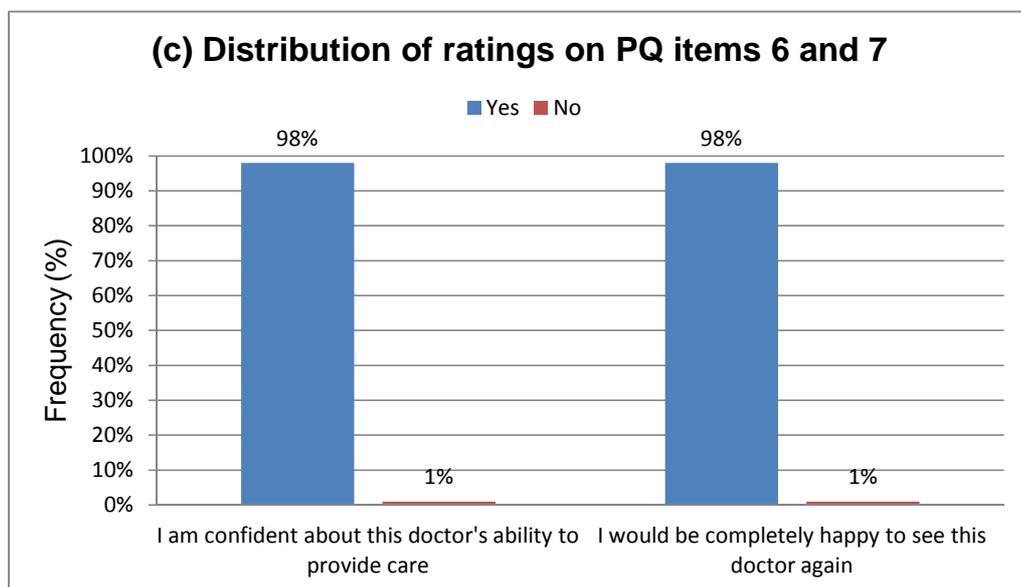
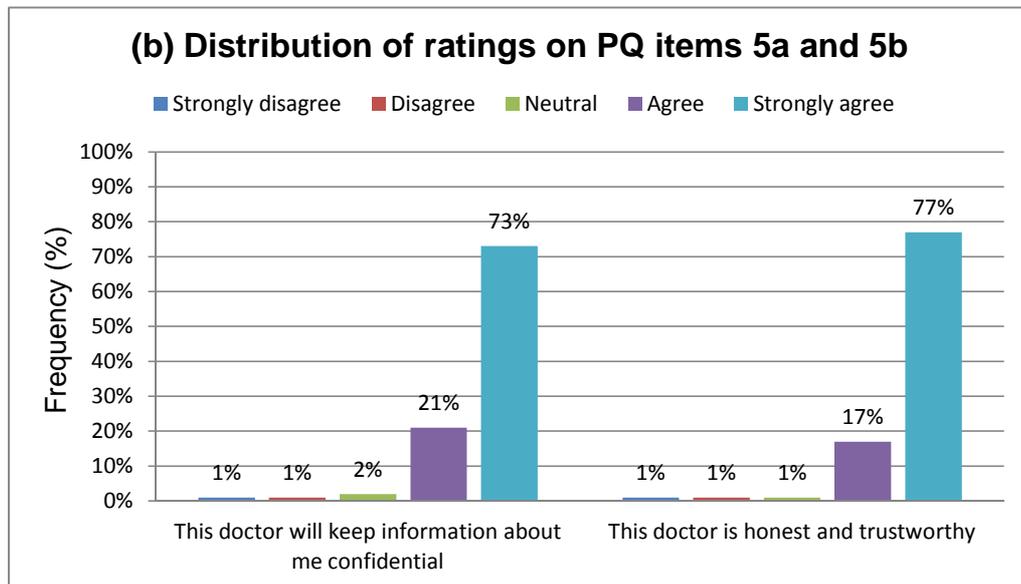
The doctor and appraiser may wish to consider:

- ✘ Are there marked differences between the distribution of the doctor's ratings and the distribution of ratings that were achieved by other doctors on the same item?
- ✘ If so, do these differences suggest particular areas of strength or weakness in the doctor's performance when compared to their peers?

**Figure 1: Patient ratings on the GMC Patient Questionnaire (PQ) – based on responses obtained from 30,333 patients in respect of 922 doctors (in recent pilot work: 2008-2010)<sup>13;14</sup>**



**Figure 1 (continued)**



### **5.3.3 Benchmark table**

The benchmark table presents: (a) the doctor's mean percentage score on each item of the PQ; and (b) the range of scores that other doctors have achieved on the same item.

It is strongly recommended that doctors and appraisers read the "Important notes about the benchmark data" that appear beneath the benchmarks tables. If necessary, they can also refer to the "Supporting documents" section of the report which explains how the mean percentage scores have been calculated and what the quartile bandings mean.

#### **(a) Checking the doctor's overall scores**

The benchmark table provides the doctor's overall (mean percentage) score on each questionnaire item, with a score of 100% representing that all patients rated the doctor's performance in that area as 'Very good'.

- ① *Because most patients have favourable views of doctors, the typical score profile is skewed towards the upper end of the percentage scale.*
- ① *It is important to recognise that, whilst the doctor's percentage scores may be high, small differences between percentage scores on individual items may reflect important distinctions in patients' perceptions of different aspects of the doctor's practice.*

The overall scores can be used to reflect further on the doctor's strengths and weakness. The doctor and the appraiser might wish to discuss:

- ✘ On which questionnaire item(s) does the doctor have the highest mean percentage score?
- ✘ On which item(s) does the doctor have the lowest mean percentage score?

### ***(b) Benchmarking the doctor's performance***

The benchmark table also allows the doctor and appraiser to compare the doctor's overall scores on the PQ items with those achieved by their peers. The table shows the range of scores that other doctors have received on each item as well as the values for the lower quartile, median and upper quartile thresholds. The doctor's mean percentage score on each item is placed into one of three colour-coded benchmark bandings, depending on how the doctor's score compares with those achieved by other doctors.

*On each item, the doctor's score will be placed in either:*

- ① The upper quartile – the doctor's overall score falls within the range of the scores achieved by the top 25% of other doctors (i.e. between the upper quartile and maximum score values);*
- ① The middle two quartiles – the doctor's overall score falls within the range of the scores achieved by the middle 50% of doctors (i.e. between the lower quartile and upper quartile values);*
- ① The lower quartile – the doctor's overall score falls within the range of the scores achieved by the lowest 25% of doctors (i.e. between the lower quartile and minimum score values).*

When reviewing the table, the doctor and appraiser may wish to discuss:

- ✘ On how many items does the doctor's overall score fall into the:
  - Upper quartile benchmark banding?
  - Middle quartile benchmark banding?
  - Lower quartile benchmark banding?
- ✘ Does the pattern of the doctor's overall scores and benchmark bandings for the different PQ items confirm the areas of potential strength and weakness identified from the frequency and distribution table?

To further understand how the doctor's performance compares with that of their peers, the appraiser may also consider where the doctor's overall score on each item lies within the assigned benchmark banding – for example:

- If the doctor is placed among the highest 25% of all doctors, is their overall score nearer to the threshold for the upper quartile or nearer to the maximum score?
- If the doctor is placed among the middle 50% of all doctors, is their overall score nearer to the threshold for the upper quartile or for the lower quartile?
- If the doctor is placed among the lowest 25% of all doctors, is their overall score nearer to the threshold for the lower quartile or nearer to (or below) the minimum score?

When reflecting on the benchmark table, the doctor and appraiser may wish to refer back to the frequency and distribution table to determine whether the doctor's reported overall score on any item could be affected by a small number of 'Satisfactory' (or lower) ratings.

The appraiser may wish to consider whether the doctor's overall MSF scores could have been affected by characteristics of the doctor or the context in which they work.

In recent research,<sup>13</sup> less favourable patient feedback was received by doctors who:

- ① *Obtained their primary medical degree from a non-European country;*
- ① *Practised as psychiatrists.*

#### **5.3.4 Free text comments**

Finally, the doctor and appraiser may review the free text comments made by patients to gain a better understanding of the doctor's overall scores. They may wish to discuss:

- ✘ Are there common themes among the comments – do a number of patients refer to the same issue?
- ✘ What do the comments suggest the doctor does well / less well?
- ✘ Do the comments indicate why the doctor received higher or lower ratings from their patients on particular items?
- ✘ Do any of the comments indicate a need for further training or action?

### **5.4 The colleague feedback section**

It is recommended that the tables in the colleague feedback section are reviewed in the order in which they are presented.

#### **5.4.1 Information about the colleague sample**

As an initial step, the doctor and appraiser may wish to reflect on the process used to nominate colleagues, the number of questionnaires returned, and the characteristics of the colleague sample. The following aspects might be discussed:

##### **(a) The number of colleagues surveyed**

- ✘ How many colleagues did the doctor nominate at the start of the process?
- ✘ How did the doctor decide which colleagues to nominate?
- ✘ What mix of colleagues did they nominate?
- ✘ Have sufficient colleague questionnaires been returned?

- ① *Doctors are asked to nominate 20 colleagues who can comment on their professional practice.*
- ① *It is recommended that they nominate a mix of 10 medical colleagues and 10 non-medical colleagues.*
- ① *It may be difficult for some groups of doctors to identify the recommended number and mix of colleagues – e.g. doctors who work in small teams or work as locums.*
- ① *15 or more completed questionnaires are required to ensure that the results are sufficiently reliable for use in formative feedback.*

**(b) The characteristics of the colleague sample**

- ✘ Is the sample representative of the colleagues that the doctor usually works with, and the colleagues who were nominated to provide feedback?
- ✘ Might the views of particular professional groups be overrepresented in the survey?

- ① *Ideally feedback should be provided by a balanced mix of medical and non-medical colleagues.*
- ① *Colleagues in managerial or administrative roles, and health professionals in non-medical roles, tend to give more favourable feedback than medical colleagues.*
- ① *Colleagues who have more frequent contact with the doctor tend to give more favourable feedback.*

The doctor and appraiser may wish to consider whether any of the above factors could have affected the results obtained in the colleague survey.

**5.4.2 Frequency and distribution table**

To obtain a balanced view of the feedback obtained, the doctor and the appraiser should review and reflect on the table that presents the frequency and distribution of colleague ratings. The following aspects might be considered:

**(a) The proportion of 'valid' responses**

- ✘ What level of missing/spoilt data does the doctor have for each item?

- ✘ What proportions of colleagues used the 'Don't know' response option for each item?

In recent pilot work:<sup>14</sup>

- ① *Missing and spoilt data was minimal (<1%) across the 19 core CQ items.*
- ① *The 'Don't know' response option varied across the CQ items (range 1% to 28% colleagues).*
- ① *Items relating to prescribing, reviewing and reflecting on one's practice, teaching, and supervision showed the highest use of the 'Don't know' response option (23% to 28% colleagues).*
- ① *The 'Don't know' response option was less frequently used (1% to 4% colleagues) on the probity and health items (Q16 to Q18).*

- ✘ On each item, how many 'valid' responses are there?

Valid responses utilise the following options on the rating scale and exclude 'Does not apply' responses, and missing or spoilt data:

- ① *'Poor' to 'Very good' (Items 1 to 15).*
- ① *'Strongly disagree' to 'Strongly agree' (Items 16 to 18).*
- ① *'Yes' or 'No' (Item 19)*

### **(b) The distribution of 'valid' responses**

For each item on the CQ, the doctor and the appraiser might wish to identify:

- ✘ The range (or spread) of responses across the scale.
- ✘ The response option(s) that colleagues have used most commonly to rate the doctor's performance.
- ✘ Whether colleague ratings are mainly positive, neutral or negative.

### **(c) Identifying possible areas of relative strength and weakness**

Having reviewed the frequency and distribution of scores, the doctor and appraiser may wish to discuss:

- ✘ Does the distribution of colleague ratings vary across the different items on the questionnaire?
- ✘ Are there any obvious areas of strength or weakness in the doctor's performance based on the colleague ratings?

The GMC's guidance for doctors<sup>4</sup> recommends that:

- ① *"The discussion at appraisal should highlight areas of good performance and help you to identify any areas that might require further development".*

It may be helpful to compare the distribution of the doctor's colleague ratings with the distribution of colleague ratings in recent pilot work<sup>14</sup> – see Figure 2.

For example, on Item 1 of the CQ ('Clinical knowledge'):

- ① *Most colleagues in the pilot work selected the 'Very good' (69%) or 'Good' (21%) response options.*
- ① *Smaller proportions of colleagues selected the 'Satisfactory' (2%) or 'Less than satisfactory' (1%) response options.*
- ① *No colleagues selected the 'Poor' response option.*

The doctor and appraiser may wish to consider:

- ✘ Are there marked differences between the distribution of the doctor's ratings and the distribution of ratings that were achieved by other doctors on the same item?
- ✘ If so, do these differences suggest particular areas of strength or weakness in the doctor's performance when compared to their peers?

**Figure 2: Colleague ratings on the GMC Colleague Questionnaire (CQ) – based on responses obtained from 17,012 colleagues in respect of 1057 doctors (in recent pilot work: 2008-2010)<sup>13;14</sup>**

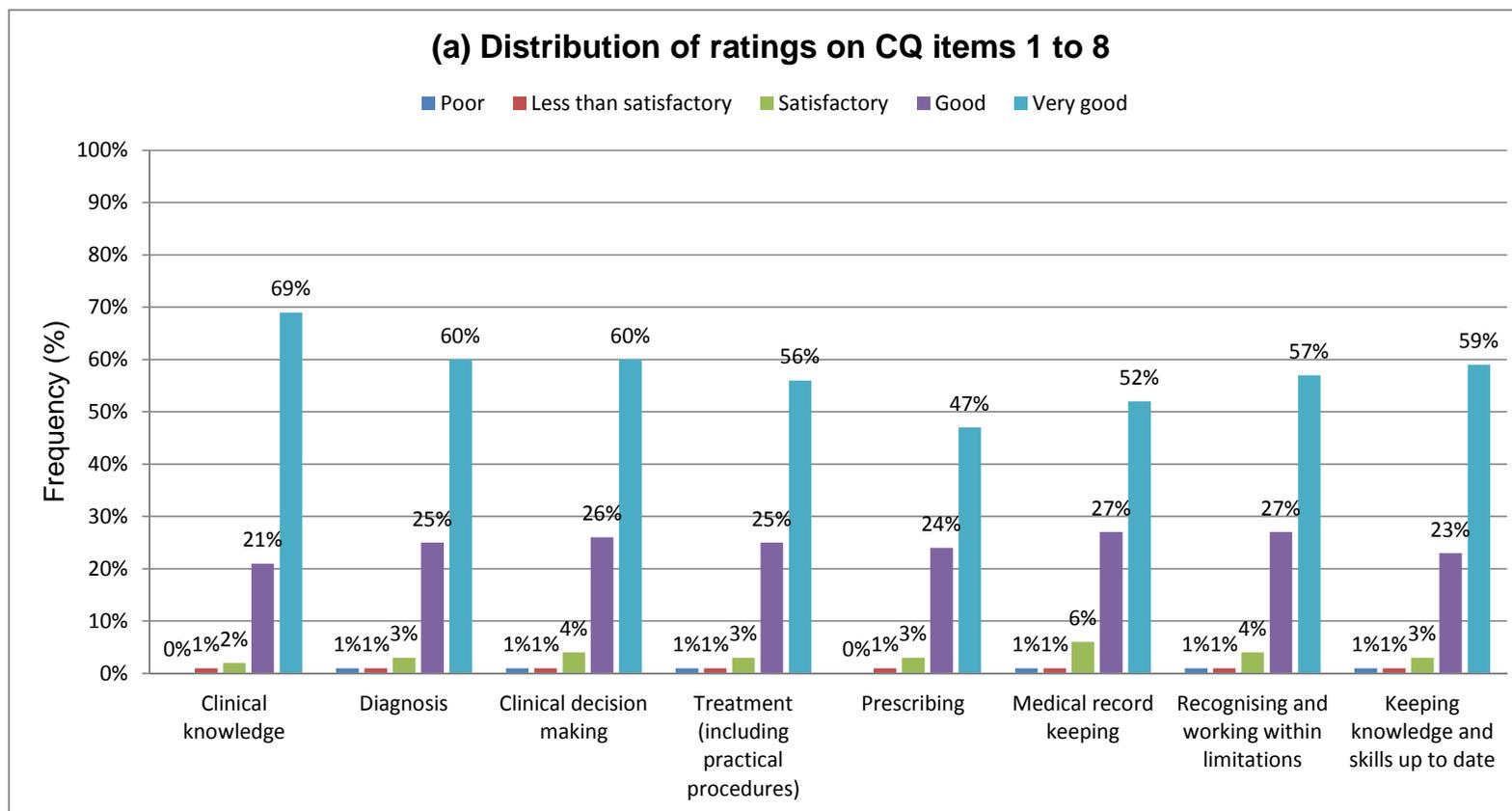


Figure 2 (continued)

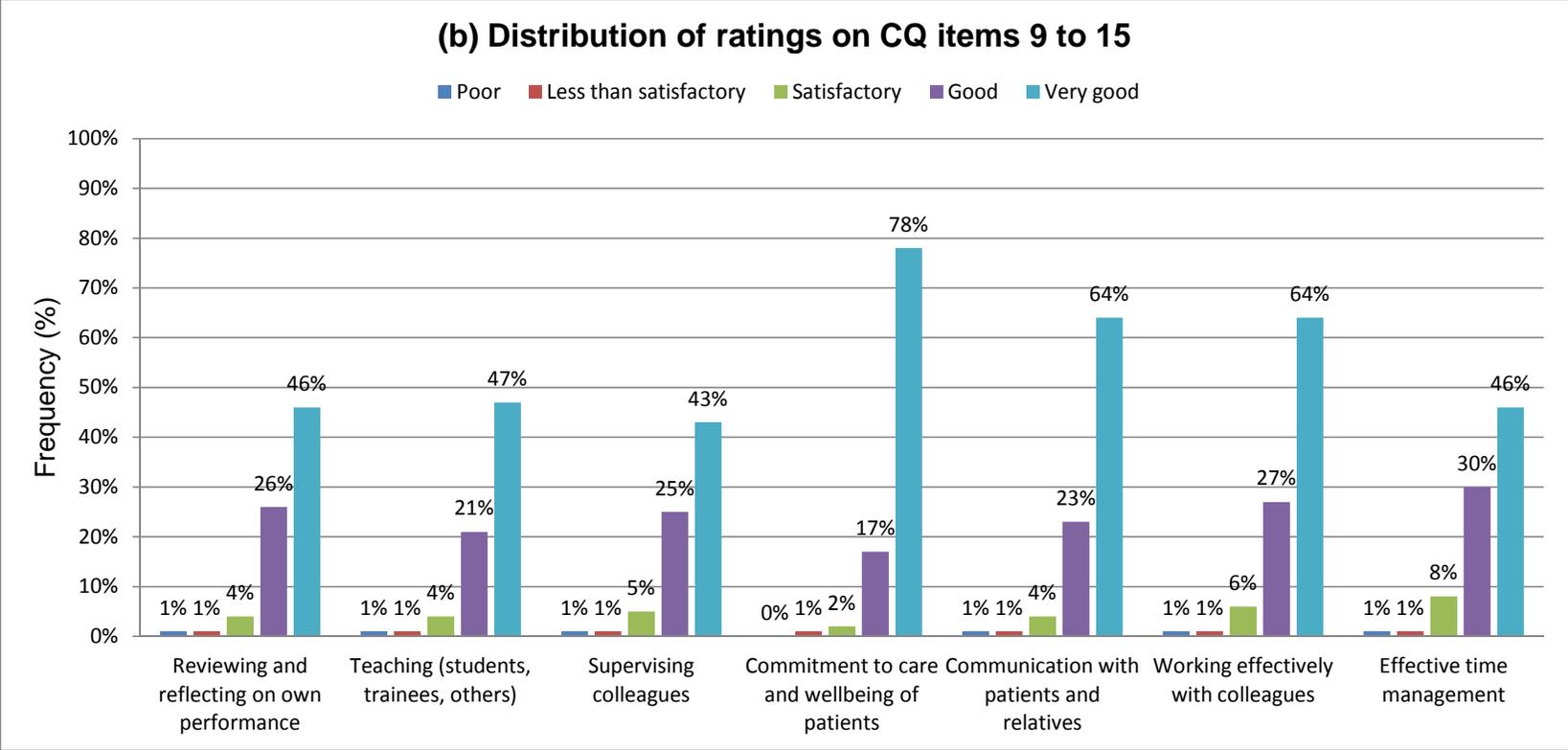
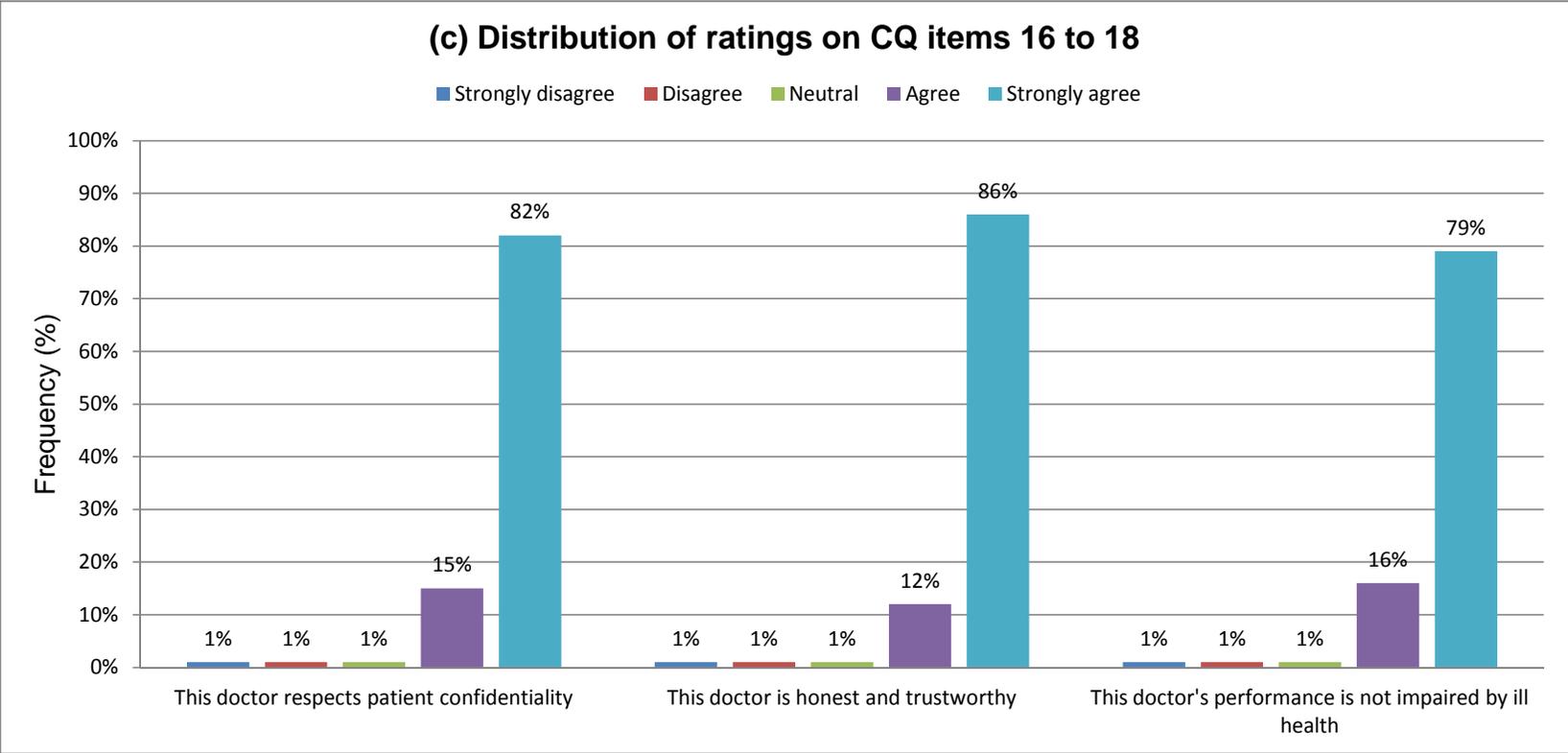
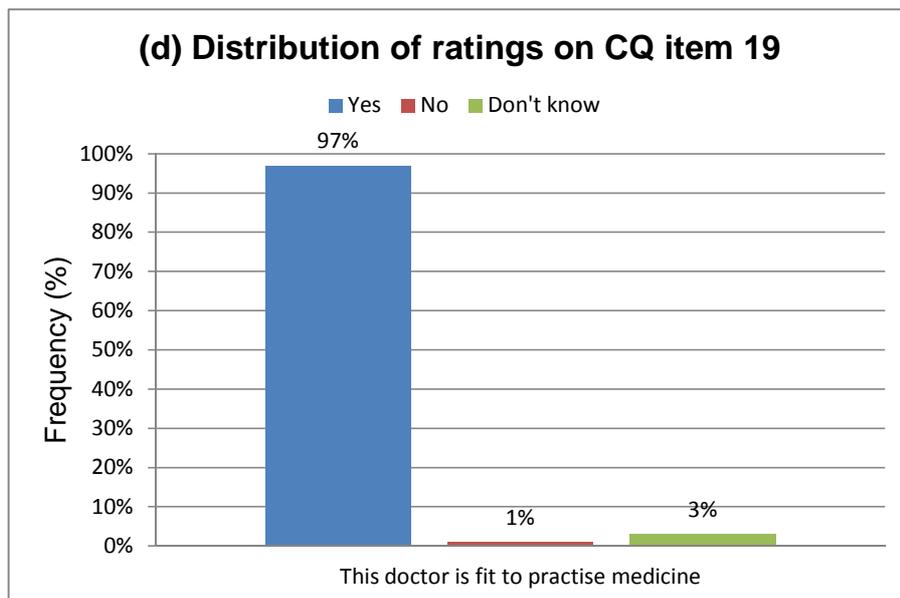


Figure 2 (continued)



**Figure 2 (continued)**



### **5.4.3 Benchmark table**

The benchmark table presents: (a) the doctor's mean percentage score on each item of the CQ; and (b) the range of scores that other doctors have achieved on the same item.

It is strongly recommended that doctors and appraisers read the "Important notes about the benchmark data" that appear beneath the benchmarks tables. If necessary, they can also refer to the "Supporting documents" section of the report which explains how the mean percentage scores have been calculated and what the quartile bandings mean.

#### **(a) Checking the doctor's overall scores**

The benchmark table provides the doctor's overall (mean percentage) score on each questionnaire item, with a score of 100% representing that all colleagues rated the doctor's performance in that area as 'Very good'.

- ① *Because most colleagues have favourable views of doctors, the typical score profile is skewed towards the upper end of the percentage scale.*
- ① *It is important to recognise that, whilst the doctor's percentage scores may be high, small differences between percentage scores on individual items may reflect important distinctions in colleagues' perceptions of different aspects of the doctor's practice.*

The overall scores can be used to reflect further on the doctor's strengths and weakness. The doctor and the appraiser might wish to discuss:

- ✕ On which questionnaire item(s) does the doctor have the highest mean percentage score?
- ✕ On which item(s) does the doctor have the lowest mean percentage score?

#### ***(b) Benchmarking the doctor's performance***

The benchmark table also allows the doctor and appraiser to compare the doctor's overall scores on the CQ items with those achieved by their peers. The table shows the range of scores that other doctors have received on each item as well as the values for the lower quartile, median and upper quartile thresholds. The doctor's mean percentage score on each CQ item is placed into one of three colour-coded benchmark bandings, depending on how the doctor's score compares with those achieved by other doctors.

*On each item, the doctor's score will be placed in either:*

- ① *The upper quartile – the doctor's overall score falls within the range of the scores achieved by the top 25% of other doctors (i.e. between the upper quartile and maximum score values);*
- ① *The middle two quartiles – the doctor's overall score falls within the range of the scores achieved by the middle 50% of doctors (i.e. between the lower quartile and upper quartile values);*
- ① *The lower quartile – the doctor's overall score falls within the range of the scores achieved by the lowest 25% of doctors (i.e. between the lower quartile and minimum score values).*

When reviewing the table, the doctor and appraiser may wish to discuss:

- ✘ On how many items does the doctor's overall score fall into the:
  - Upper quartile benchmark banding?
  - Middle quartile benchmark banding?
  - Lower quartile benchmark banding?
  
- ✘ Does the pattern of the doctor's overall scores and benchmark bandings for the different CQ items confirm the areas of potential strength and weakness identified from the frequency and distribution table?

To further understand how the doctor's performance compares with that of their peers, the appraiser may also consider where the doctor's overall score on each item lies within the assigned benchmark banding – for example:

- If the doctor is placed among the highest 25% of all doctors, is their overall score nearer to the threshold for the upper quartile or nearer to the maximum score?
- If the doctor is placed among the middle 50% of all doctors, is their overall score nearer to the threshold for the upper quartile or for the lower quartile?
- If the doctor is placed among the lowest 25% of all doctors, is their overall score nearer to the threshold for the lower quartile or nearer to (or below) the minimum score?

When reflecting on the benchmark table, the doctor and appraiser may wish to refer back to the frequency and distribution table to determine whether the doctor's reported overall score on any item could be affected by a small number of 'Satisfactory' (or lower) ratings.

The doctor may wish to consider whether the doctor's MSF scores could have been affected by characteristics of the doctor or the context in which they work.

In recent research,<sup>13</sup> less favourable colleague feedback was received by doctors who:

- ① *Obtained their primary medical degree outside of the UK or South Asia;*
- ① *Practised as a general practitioner or psychiatrist;*
- ① *Were employed in a contractual role (grade) other than a consultant or general practitioner;*
- ① *Worked in a locum capacity.*

#### **5.4.4 Free text comments**

Finally, the doctor and appraiser may review the free text comments made by colleagues to gain a better understanding of the doctor's overall scores. They may wish to discuss:

- ✘ Are there common themes among the comments – do a number of colleagues refer to the same issue?
- ✘ What do the comments suggest the doctor does well / less well?
- ✘ Do the comments indicate why the doctor may have received higher or lower ratings from their colleagues on particular items?
- ✘ Do any of the comments indicate a need for further training or action?

#### **5.5 Self-assessment section**

The doctor and appraiser may wish to use this section of the report to assess the accuracy of the doctor's own perceptions about their performance – i.e. whether the areas of strength and weakness that the doctor perceives within their own performance match those identified by their patients and colleagues.

The doctor and the appraiser might wish to discuss:

- ✘ To what extent do the doctor's self-ratings match the average ratings provided by their patients and their colleagues?

- ✘ Were there any items where the doctor's self-rating was unexpectedly higher than the average rating given by their patients or colleagues (i.e. areas of relative 'weakness' they may have been unaware of)?
  
- ✘ Were there any items where the doctor's self-rating was unexpectedly lower than the average rating given by their patients or colleagues (i.e. areas of relative 'strength' they may have been unaware of)?

---

## **6 Determining what further action is needed**

Having reviewed and reflected upon each of the feedback in the doctor's personalised report, the doctor and appraiser should consider whether any further action is required and, if so, what that action should be.

### **6.1 Making judgements about the need for further action**

It may be difficult for the appraiser to make judgements about the need for further action based on a doctor's MSF report and caution should be exercised before recommending any action. Decisions should be made on an individual basis, taking into account characteristics of the patient and colleague survey samples, characteristics of the doctor and the context in which they practice medicine.

It is important for both the doctor and the appraiser to recognise that MSF is just one element of a range of evidence that doctors are expected to collect about their professional practice for the purposes of revalidation. Thus the MSF results should be considered alongside any other evidence that the doctor has accumulated in the current revalidation cycle.

### **6.2 Managing discussions about further action**

In managing the discussion about further action, the appraiser should recognise that many doctors have high expectations of themselves and some may be disappointed if their MSF scores are lower than they anticipated or particular free text comments appear unfairly critical. Appraisers may wish to reinforce the notion that no doctor can expect to be perfect in every aspect of their practice.

Appraisers may also wish to emphasise that the feedback report summarises the perceptions of a sub-sample of patients and colleagues who have been asked to rate very specific aspects of the doctor's practice. As such, the doctor can view the feedback as a tool for reflecting on their practice and, if necessary, use it to identify ways in which they might improve the way they work. The doctor should be encouraged to incorporate the feedback in the report into their future development plan, alongside their own priorities.

As well as identifying potential weaknesses or problems, the appraiser should help the doctor to recognise their strengths and to value their achievements with regard to those areas of practice.

### **6.3 Deciding on further action: what about 'low performers?'**

In deciding whether further action is indicated, the appraiser and the doctor may first wish to summarise the key messages that arise from the feedback report.

In doing so, they may wish to discuss:

- ✘ What areas of practice were highlighted as the doctor's main strengths in the patient or colleague feedback sections?
- ✘ What areas of practice were highlighted as the doctor's main weaknesses in the patient or colleague feedback sections?
- ✘ If the doctor completed a patient survey and a colleague survey, did similar messages arise from both surveys?
- ✘ In their self-assessment or individual reflection, does the doctor agree with the areas of strength and weakness identified from the survey results?

The appraiser and doctor may wish to identify specific items on the PQ or CQ where the doctor's scores were lowest, and consider whether these scores

were in fact *abnormally* low. In doing so, the doctor and appraiser should bear in mind the skewed nature of the GMC benchmark data and the rating biases that are known to be associated with certain characteristics of the doctor and their patient and colleague assessors (see Chapter 4). The doctor and appraiser may wish to discuss:

- ✘ On which questionnaire items did the doctor achieve the poorest ratings and/or the lowest overall score?
- ✘ On these items, when compared to other doctors, did the doctor fall into the lowest quartile benchmark banding and, if so, how close was the doctor's score to the minimum score provided in the benchmark data?
- ✘ Did the free text comments of patients or colleagues also highlight specific problems in these areas of practice?
- ✘ On these items, were there marked differences between the doctor's self-assessment score and the way in which patients or colleagues rated the doctor's performance?
- ✘ How do the MSF results compare to other evidence the doctor has collected about their performance?
  - For example, are the MSF results supported by audit data, clinical outcome reviews, patient letters of complaint or compliment, or informal feedback from patients or colleagues?

If several areas of 'weakness' are identified, the appraiser and doctor may wish to prioritise one or two areas of practice, and tailor the doctor's action plan towards those areas first. Having agreed the areas to prioritise, the doctor and appraiser may wish to:

- ✘ Discuss the feasibility of changing the doctor's practice in that area and establish the time/cost implications of the required action;
- ✘ Set specific and achievable goals in relation to the required change(s);

- ✘ Determine what educational support and resources are available to help the doctor achieve their goals. These might include mentoring opportunities, self-directed learning activities, online training, face-to-face training, attending local seminars.
- ✘ Consider whether there are any procedural or organisational changes or health resources that might help the doctor achieve their goals;
- ✘ Agree a plan for monitoring the doctor's progress towards their goals.

The appraiser should also consider whether it would be appropriate for the doctor to repeat the patient and/or colleague survey at a later date (for example, within a year) and, if serious concerns have been highlighted, whether it would be appropriate to discuss the doctor's feedback with the Responsible Officer or other senior colleagues within the organisation.

#### **6.4 Serious concerns**

Amongst the sample of 1,065 doctors, and 30,333 of their patients and 17,012 of their colleagues whose data formed the basis of this guidance, almost no serious concerns were raised regarding a doctor's professional practice in the course of the MSF process.

However, it is not impossible that such a concern might be raised during the course of patient and/or colleague feedback. Serious concerns might encompass such issues as sexual harassment or seriously inappropriate behaviour, or repeated/serious allegations of substance misuse, criminal behaviour, or bullying. These may be raised in the free text comments on the questionnaires.

Appraisal processes thus need to be reviewed to ensure that an appropriate and timely level of response is provided in the unlikely occurrence of this

eventuality. An 'appropriate' response needs to take account of the guidance provided here, and a 'timely' response needs to take patient safety into consideration. Appraisers (and others) will wish to note that the data obtained through patient and colleague feedback is anonymous. It is the responsibility of local systems and clinical managers/appraisal leads to ensure that robust and defensible appraisal systems are in operation locally.

---

## 7 References

1. General Medical Council. Revalidation: *The way ahead - Consultation Document*. London: General Medical Council, 2010
2. General Medical Council. Revalidation: *The way ahead - Response to our revalidation consultation*. London: General Medical Council, 2011.
3. General Medical Council. *The Good Medical Practice Framework for appraisal and revalidation*. London: General Medical Council, 2011.
4. General Medical Council. *Supporting information for appraisal and revalidation*. London: General Medical Council, 2011.
5. General Medical Council. Revalidation: *The way ahead. Annex 3 - GMC principles, criteria and key indicators for colleague and patient questionnaires in revalidation*. London: General Medical Council, 2010.
6. Chisholm A, Askham K, Picker Institute Europe. *What do you think of your doctor? A review of questionnaires for gathering patients' feedback on their doctor*. Oxford: Picker Institute Europe, 2006.
7. Evans R, Elwyn G, Edwards A. Review of instruments for peer assessment of physicians. *BMJ*, 2004; **328**(7450): 1240-1245.
8. General Medical Council. *Good medical practice: Guidance for doctors*. London: General Medical Council, 2006.
9. Campbell JL, Richards SH, Dickens A, et al. Assessing the professional performance of UK doctors: An evaluation of the utility of the General Medical Council patient and colleague questionnaires. *Qual Saf Health Care*, 2008; **17**: 187-193.
10. Kilminster S, Pell G, Roberts T. *Patient and colleague questionnaires: Validation report to the GMC*. Leeds: University of Leeds, Medical Education Unit, 2005.
11. MORI Social Research Institute. *Revalidation questionnaires testing: Qualitative research findings*. London: General Medical Council, 2004.

12. Lockyer J, Fidler H. *Comparison of colleague and patient multisource feedback instruments designed for GPs in UK*. London: Royal College of General Practitioners, 2010.
13. Campbell J, Roberts M, Wright C, Hill J, Greco M, Taylor M, Richards S. Factors associated with variability in the assessment of UK doctors' professionalism: an analysis of survey results. *BMJ*, 2011; in press.
14. Wright C, Richards SH, Hill J, Roberts M, Norman GR, Greco M, Taylor MRS, Campbell JL. Assessing professionalism: properties of the GMC Patient and Colleague Questionnaires. *Submitted*, August 2011.
15. Hill J, Richards S, Asprey A, Campbell JL. A qualitative study of the use of multisource feedback in appraisal and for revalidation. *Submitted*, October 2011.
16. Roberts M, Campbell JL, Richards SH, Wright C. Multisource feedback for revalidation: Do doctors see themselves as others do? *In preparation*, October 2011.
17. Wright C, Asprey A, Campbell JL. Obtaining patient and colleague feedback on professional performance: Is multisource feedback feasible for anaesthetists? *In preparation*, October 2011.
18. General Medical Council. *Annex 3: GMC principles, criteria and key indicators for colleague and patient questionnaires in revalidation*. London: UK General Medical Council, 2011.
19. Eva KW, Regehr G. Self-assessment in the health professions: A reformulation and research agenda. *Acad Med*, 2005; **80**: S46-S54.
20. Sargeant J. Toward a common understanding of self-assessment. *J Contin Educ Health Prof*, 2008; **28**: 1-4.
21. Archer J. State of the science in health professional education: effective feedback. *Med Educ*, 2010; **44**: 101-108.
22. Lockyer JM, Violato C, Fidler H. Likelihood of change: a study assessing surgeon use of multisource feedback data. *Teaching and Learning in Medicine*, 2003; **15**: 168-174.

23. Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness. *Med Educ*, 2005; **39**: 497-504.
24. Sargeant JM, Mann K, van der Vleuten CPM, Metsemakers J. 'Directed' self-assessment: practice and feedback within a social context. *J Contin Educ Health Prof*, 2008; **28**: 47-54.
25. Hodges B, Reger G, Martin D. Difficulties in recognising one's own incompetence: novice physicians who are unskilled and unaware of it. *Acad Med*, 2001; **76**(10): S87-S89.
26. Davis DA, Mazmanian PE, Fordis M, van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence. *JAMA*, 2006; **296**: 1094-1102.
27. Wright C on behalf of the GMC research team. *GMC Multi-Source Feedback Study. Nested Study III: Development of personalised feedback reports for doctors*. Exeter, UK: Peninsula College of Medicine and Dentistry, March 2011.
28. Richards S, Campbell J, Walshaw E, Dickens A, Greco M. A multi-method analysis of free-text comments from the UK General Medical Council Colleague Questionnaires. *Med Educ*, 2009; **43**: 757-766.
29. Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof*, 2003; **23**(1): 4-12.

---

## Appendix 1: Content of the GMC Patient Questionnaire

### (A) Core items

Item		Response scale
4a	Being polite	Evaluative <sup>a,b</sup>
4b	Making you feel at ease in his / her presence	Evaluative
4c	Listening to you	Evaluative
4d	Assessing your condition	Evaluative
4e	Explaining your condition and treatment	Evaluative
4f	Involving you in decision about your treatment	Evaluative
4g	Providing or arranging treatment for your	Evaluative
5a	This doctor will keep information about me confidential	Agreement <sup>b,c</sup>
5b	This doctor is honest and trustworthy	Agreement
<b>Summative items</b>		
Q6	I am confident about this doctor's ability to provide care	Yes/No
Q7	I would be completely happy to see this doctor again	Yes/No

<sup>a</sup> Valid evaluative response categories (scale number): 'Poor' (1), 'Less than satisfactory' (2), 'Satisfactory' (3), 'Good' (4), 'Very good' (5).

<sup>b</sup> 'Does not apply' response option also available.

<sup>c</sup> Valid agreement response categories (scale number): 'Strongly disagree' (1), 'Disagree' (2), 'Neutral' (3), 'Agree' (4), 'Strongly Agree' (5).

## Content of the GMC Patient Questionnaire (continued)

### (B) Other items

Item		Response categories
1	Who is filling in the patient questionnaire?	Patient Parent Spouse/partner Relative/friend
2	Reason(s) why the patient saw the doctor*	To ask for advice Because of an ongoing problem For treatment (inc. prescriptions) Because of a one-off problem For a routine check For other reason
3	How important the visit was to patient's health/wellbeing?	5 point scale, anchored at 'Not very important' (1) and 'Very important' (5).
8	Was the patient's visit with their usual doctor?	Yes/No
9	<i>Please add any other comments you have about this doctor</i>	Free text box
10	Gender	Male/Female
11	Age	Under 15 years 15 to 20 years 21 to 40 years 41 to 60 years Over 60 years
12	Ethnicity	White Mixed Asian or Asian British Black or Black British Chinese/Other

\* Respondents can tick more than one response option.

---

## Appendix 2: Content of the GMC Colleague Questionnaire

<b>(A) Core items</b>		<b>Response scale</b>
<b>Item</b>		
1	Clinical knowledge	Evaluative <sup>a,b</sup>
2	Diagnosis	Evaluative
3	Clinical decision making	Evaluative
4	Treatment (including practical procedures)	Evaluative
5	Prescribing	Evaluative
6	Medical record keeping	Evaluative
7	Recognising and working within limitations	Evaluative
8	Keeping knowledge and skills up to date	Evaluative
9	Reviewing and reflecting on own performance	Evaluative
10	Teaching (students, trainees, others)	Evaluative
11	Supervising colleagues	Evaluative
12	Commitment to care and wellbeing of patients	Evaluative
13	Communication with patients and relatives	Evaluative
14	Working effectively with colleagues	Evaluative
15	Effective time management	Evaluative
16	This doctor respects patient confidentiality	Agreement <sup>b,c</sup>
17	This doctor is honest and trustworthy	Agreement
18	This doctor's performance is not impaired by ill health	Agreement
19	This doctor is fit to practise medicine	Yes/No <sup>b</sup>

<sup>a</sup> Valid evaluative response categories (scale number): 'Poor' (1), 'Less than satisfactory' (2), 'Satisfactory' (3), 'Good' (4), 'Very good' (5).

<sup>b</sup> Don't know' response option also available.

<sup>c</sup> Valid agreement response categories (scale number): 'Strongly disagree' (1), 'Disagree' (2), 'Neutral' (3), 'Agree' (4), 'Strongly Agree' (5).

## Content of the GMC Colleague Questionnaire (continued)

### (B) Other items

Item		Response categories
20	<i>Please add any other comments you want to make about this doctor.</i>	Free text box
21	Gender	Male/Female
22	Age group	16 to 19 years 20 to 29 years 30 to 39 years 40 to 49 years 50 to 59 years 60 years or over
23	Professional role	Doctor Registered Nurse Health Visitor/Midwife Pharmacist Administrator/Receptionist/Secretary Allied Healthcare Professional Health Care Assistant Non-clinical Manager Other professional role
24	How recently familiar with the doctor's clinical practice	Current colleague Within last 2 years Between 2 and 5 years Between 5 and 10 years More than 10 years ago
25	Frequency of contact with the doctor	Most days Weekly Monthly Less often
26	Ethnic group	White Mixed Asian or Asian British Black or Black British Chinese/Other