

RACHEL LEA *Practice Development Manager, Beds & Herts LMC*

MICHAEL HARRISON *Executive Director, Beds & Herts LMC*



Bedfordshire & Hertfordshire
Local Medical Committee Ltd

PRACTICE HEALTHCHECK
QUESTIONNAIRE (PHQ)

ANALYSIS & REVIEW

2019

In partnership with



Berkshire, Buckinghamshire
& Oxfordshire LMCs



Executive Summary

Background

Over recent years there has been a growing focus on GP practice resilience, driven primarily by increasing numbers of practices either handing back their contracts, or having to merge with another practice in the hope of finding stability.

Often when organisations talk about the assessment of practice resilience, the term “resilience” gets conflated with other more measurable practice attributes, such as clinical quality and performance metrics. As a result, true practice resilience is poorly understood as it is difficult to objectively measure.

The resilience work undertaken by Bedfordshire and Hertfordshire LMC in 2017 as part of the General Practice Forward View (GPFV) showed that a short practice health check questionnaire (PHQ) was a good surrogate measure for overall practice resilience as the outcomes strongly correlated with the outcomes of an in-depth resilience assessment.

As a result of this initial work, BLMK STP and East & North Herts CCG provided funding for the LMC to redevelop the PHQ into a web-based tool, based on the learning from the original programme. The revised PHQ used 23 metrics grouped into five sections: workforce, infrastructure, internal communications, workload pressures and relationships.

The PHQ was sent out to all GMS practices in Bedfordshire, Luton, Hertfordshire and Milton Keynes. Each practice that completed the PHQ received a two-page practice specific feedback report, showing the overall practice resilience score, their rating on each of the 23 metrics, and how their data compared to other local practices.



Data Analysis

In total, 138 practices completed the PHQ and received an individualised feedback report. The overall response rate across the five CCGs was 70%. The average practice resilience score was 73%, the lowest score being 44% and the highest being 97%. Practices which scored below 60% were seen as potentially vulnerable and those that scored above 80% were seen as resilient.

The combined dataset of 138 practices was then analysed to identify the factors that had a statistically significant impact on practice resilience (using a two-tailed T-test).

Firstly, factors related to the size and structure of the practice were analysed. For example, analysis of practice actual list size showed a statistically significant difference in resilience score (practices with a list size of 9,000 or above had a resilience score that was on average 6 points higher than those below 9,000).

Secondly, the 23 metrics used to calculate the overall resilience score were analysed. Fourteen showed a statistically significant difference in practice resilience score. They were:

1. GPs expected to leave over the next year
2. Locum usage
3. Current GP vacancies
4. Management of home visit requests
5. Recruitment of GPs
6. Contractor clinical hours in the practice
7. Partnership agreements
8. Impact of losing one full-time equivalent contractor
9. List closure requests
10. Change in contractor drawings
11. Premises issues
12. Regularity of partner meetings
13. Relationship with neighbouring practices
14. Training practice status

Details of the impact these metrics had on resilience score are given in section 3.

While the individualised practice data has not been shared with anyone beyond the practice itself, aggregated data at CCG level has been used to provide feedback to the commissioners on how they can better support practices to build resilience.

The data showed significant variation between CCGs. For example, when analysing the number of practices that had more than 10% of the GP sessions vacant (a metric which proved to be statistically significant) we found that in one CCG only 12% of practices fell into this category, compared to 52% in another CCG. Examples like this showed the variation in the challenges that practices face across the region, and therefore the need for a more localised approach to delivering support to practices.

The final area of analysis was focused on sub-dividing the practice resilience score into two separate scores, called the predictive and symptomatic scores. The predictive score looked at factors that affect the underlying stability of the practice, such as if they have a partnership agreement in place. The symptomatic score looked at factors that were symptomatic of low resilience, rather than a cause, such as the reliance on locums.



Impact

The tool has proved helpful to practices as a way to gain a snapshot of their resilience levels and identify the areas they need to focus on to try and build for the future.

From an LMC perspective it has helped to develop a much broader overview of the resilience of practices within our area and has enabled us to reach out to practices with the offer of support that would not otherwise have been on our radar.

At a commissioner level it has helped the CCG gain an insight into some of the resilience challenges faced by their practices and identified trends in the data of which they may not have otherwise been aware.

Future Developments

While the cycle of redesign, data collection, practice feedback, data analysis and reporting has taken 18 months to complete, this does not signify the end of the project. Following the success of the work completed to date, we plan to take the following steps in the coming months to continue the development of the PHQ tool.

- Improve both the PHQ and practice feedback reports, based on the learning from this version, and go back out to practices in 2020.
- Develop our analysis into the predictive and symptomatic resilience metrics which may help better tailoring of the support offered to practices by the LMC or CCGs.
- Continue to work with the practices identified as at risk by the PHQ to ensure appropriate support is available to them.
- Continue to work with the local CCGs and STPs to help inform how funding that is available to support practice resilience can be used in the most efficient way to benefit the practices that need it most.



Introduction

The Practice Healthcheck Questionnaire (PHQ) was designed to give an overview of a practice's current state of resilience. It was based on previous resilience work carried out by Beds & Herts LMC that had identified a number of factors that contribute to whether a practice could be defined as fragile, sustainable or thriving.

The PHQ included questions about the size and structure of the practice, workforce, workload, premises and relationships with other practices. The answers were used to create 23 metrics which were combined to give an overall practice resilience score.

The questionnaire was distributed to all GMS practices in Bedfordshire, Luton, Milton Keynes and Hertfordshire between September 2018 and January 2019. 138 practices (out of 196 GMS practices in total) completed the questionnaire, giving a completion rate of 70%.

This report is in five sections:

- Section 1** describes the background to the PHQ
- Section 2** describes how the metrics were developed and analysed
- Section 3** reports on the overall results
- Section 4** reports on the results at CCG level
- Section 5** reports on our next steps planned for 2020/21



Section 1: Background

Over the years, there have been many different attempts to measure general practice through the use of “dashboards” or “scorecards”. These largely focused on areas of quality, activity or performance, such as numbers of complaints, achievement against QOF indicators, numbers of referrals to secondary care etc, and have been used by the commissioners to compare and rank practices. The definition of a “good” practice was therefore one that met a number of externally determined criteria and “poorly performing practices” could expect to be visited by the commissioners on a regular basis.

In 2016 NHS England published the GP Forward View (GPFV) in response to an alarming increase in practices getting into difficulties and handing back their contracts. The GPFV recognised that general practice was struggling with the pressures of reduced GP workforce, increased workload, increased patient need and demand, and premises problems. This reflected the local experience where Beds & Herts LMC was starting to see practices becoming unviable and handing their contracts back, and sometimes these were practices that had met the commissioners’ definition of “good”. The LMC was keen to develop a different way to measure practices as businesses to help identify at an early stage practices that may need some support to prevent a contract termination, as well as to help understand what makes some practices more resilient and able to thrive in difficult times than others.

Part of the GPFV allocated money to local NHS England teams for “resilience”, that is, for interventions that could prevent a practice from handing its contract back. At that time, there were few options available for commissioners and the local NHS England team agreed to fund Beds & Herts LMC to develop a “resilience” package which could be offered to practices that applied for support under the GPFV. The LMC’s resilience package was taken up by about 40 practices in 2017 and 2018. It comprised a short questionnaire which was then followed up with a more detailed analysis of a practice’s situation. From the information gathered, it was clear that there was a correlation between a practice’s resilience score from the short questionnaire and its overall situation when examined in



more detail through the fuller assessment tool. Therefore, once the first wave of the resilience project had been completed in 2018, we turned our attention back to the short questionnaire, refining it and improving it, to create the Practice Healthcheck Questionnaire (PHQ). This was sent to all practices in Bedfordshire and Luton in September 2018 and to all practices in Hertfordshire in January 2019. Working in collaboration with Dr Matt Mayer at Berkshire, Buckinghamshire & Oxford (BBO) LMC, practices in Milton Keynes were also sent the PHQ, as Milton Keynes forms part of the BLMK STP.

The PHQ was designed for practices working under a partnership or as a single-handed contractor under a traditional GMS contract (or PMS, although there are no longer any PMS contracts in Bedfordshire, Luton, Milton Keynes or Hertfordshire). Some of the questions related to the clinical input of the contract holders, the profits/drawings, and – for partnerships – how the partnership functioned. These questions were integral to the calculation of the resilience score. The questionnaire was not designed for APMS contracts run by a separate organisation, the directors of which may or may not have any direct clinical input into the practice. By mistake, two APMS practices in Bedfordshire did complete the questionnaire, and it was clear that their contractual arrangements meant that some of the questions were not relevant and the calculation of the metrics did not work. These questionnaires have been excluded from the analysis.

It is important to stress that the resilience score of a practice cannot and must not be used as an indicator of a practice's performance or quality of care. We have kept all practice information confidential within the LMC and have not shared it with any other practices or any other organisation (including the CCGs) except with the explicit consent of the practice. Where a practice has raised concerns about their situation, or the PHQ has identified a practice that is vulnerable, we have encouraged the practice to approach their CCG for support and suggested that they share their PHQ results.

The design, distribution and analysis of the PHQ was kindly funded by BLMK STP and East & North Hertfordshire CCG. More detailed presentations of the results for these areas have been given to the relevant CCGs.



Section 2: Process

The primary intention of the questionnaire was to provide each practice with an overall resilience score. This was to enable the practice to assess where it lies on the scale from fragile to thriving, and also to enable the LMC to identify any practices that may be particularly vulnerable and need support. A secondary intention was to use the data to further validate the metrics and identify those elements that had a significant impact on a practice's resilience.

The questions on the PHQ were based on those previously used in the earlier resilience work which had, in turn, been based on work carried out by Devon LMC. We added some additional questions that had emerged through the 2017/18 resilience work as being relevant. Much of the information had not been collected before in this way and there were few benchmarks or standards to measure against, so we used a number of methods to create a scoring mechanism which are outlined below.

The PHQ contained between 28 and 58 questions about the practice structure, infrastructure and clinical workforce. The PHQ was an on-line interactive questionnaire and the actual number of questions any practice was asked depended on their structure and the clinical staff they employed. The answers to these questions were then used to create 23 metrics which were individually scored to produce the overall resilience score. The 23 metrics were grouped into five sections: workforce, infrastructure, internal communications, workload pressures and relationships. Appendix A lists all the metrics along with their rationale.

2.1 Weighting

When we produced the PHQ we did not know whether the metrics we were using were of equal weight: for example, does the number of GP hours per 1,000 patients per week have more, less or equal impact on a practice's resilience than whether the number of complaints has increased or

decreased, or whether or not there is an up to date partnership agreement in place? Through discussions with practice managers and GPs, and by reference to the previous resilience work, we established three different levels of importance for the metrics (high, medium and low) in terms of the impact we thought that metric would have on the practice's resilience. The list of metrics in appendix A includes the weighting we gave each metric.

2.2 Scoring

Most of the metrics were based on questions which had fixed answers (i.e. yes/no or a limited number of answer options). For all of these, one of the answer options generated the highest (best) score with the other possible answers being scored lower. Weighting of the metrics (as described in the paragraph above) meant that the possible score for each metric varied.

A number of metrics were based on numerical information from questions about number of hours or sessions of various staff groups, use of locums, vacancies and expected future losses. We used this information to create metrics such as hours per 1,000 patients per week of current staff, or percentage of current GP time the practice is expecting to lose over the next year. For some of these metrics we could use the previous resilience work to help set the scoring, but for others we had to use comparison against the average, or through consensus. We made some assumptions to set the scores for these metrics, such as “the more hours per 1,000 patients per week the better”. The results reported in the next section (Section 3) are based on this scoring system.

2.3 Feedback to Practices

Each practice received a two-page report showing the practice overall resilience score, and how this compared with other practices, and more detailed information about how the practice scored (in terms of a RAG rating) for each individual metric. All practices received their feedback report by April 2019. Figure 1 below shows the information contained on the feedback reports.

Figure 1: Practice feedback report



2.4 Further analysis of the data

As the primary intention of the PHQ was to provide feedback for each practice, our first priority was to present the results back to practices. This was completed by April 2019. Once we had completed

the practice feedback, we were able to run further analysis on the data to test which of the metrics were statistically significant and, as a result of this, whether any of the scoring needed adjusting for future use.

To test for statistical significance, we systematically removed the metric we were testing to ensure it did not bias the results. We then tested the results using a two-tailed T-test, with a p-value of 0.05 (a confidence level of 95%) to see if the difference between the two groups we were comparing was statistically significant.

An example of the process to test for statistical significance

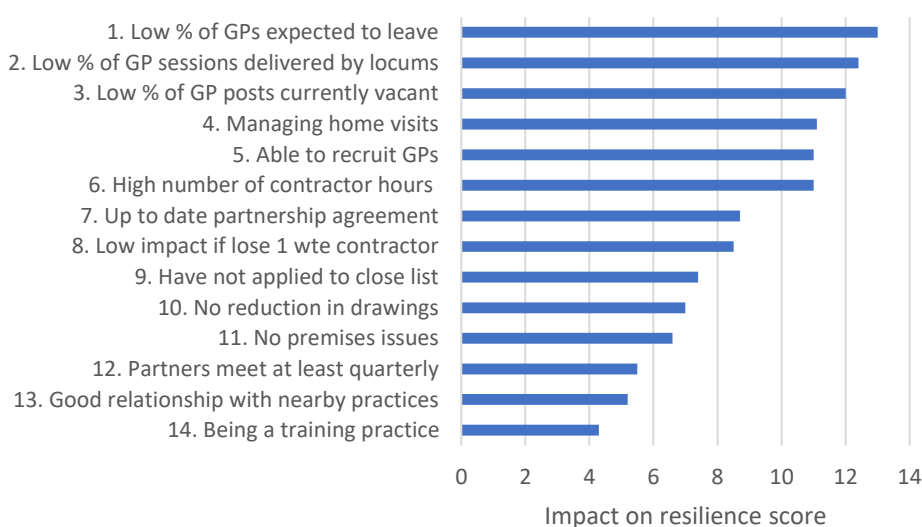
The average resilience score for training practices was 76% compared to 68% for non-training practices. However, training practice status was heavily weighted within our scoring system and so we would expect training practices' resilience scores to be higher for this reason alone. Therefore, in order to test if the difference in resilience score for the two groups was statistically significant, we first needed to remove that metric from the scoring system and then recalculate the score for each practice without the score for training practice status.

The recalculated resilience scores showed that even without the score for training practice status, training practices had a higher average score than non-training practices (77% and 73% respectively). We then ran a two-tailed T-test to check if this difference was statistically significant (at a confidence level of 95%), which it was.

We repeated this process with each metric in turn. In this way, 14 of the 23 metrics were shown to be statistically significant, and 9 were not.

2.5 Statistically significant metrics

The graph below shows those metrics that were statistically significant, and the impact each metric had on overall resilience score. Further detail about each of these 14 metrics is given in the next section.





Section 3: Overall Results

This section reports on the overall results from the 138 practices that completed the PHQ. Section 4 presents the results by CCG.

3.1 Overall results

Practice overall resilience scores ranged from 44% to 97% with an average of 73%. We considered a score below 60% as indicating a practice that could be considered vulnerable and a score of above 80% as indicating a practice that could be considered resilient. Those between 60% and 80% we considered to be stable. Out of the 138 practices

- 23 were vulnerable
- 74 were stable
- 41 were resilient

Practices were asked to rate themselves as to whether they felt fragile, sustainable or thriving. Out of the 138 practices

- 29 considered themselves to be fragile
- 91 sustainable
- 18 thriving

The table overleaf compares the practices' self-assessment with their resilience score and identifies 12 practices that were considered both vulnerable according to their score and fragile according to their self-identification. Of these, the LMC is already working with (or has worked with) six and has offered support to four (which has not yet been taken up). An offer of support for the remaining two practices is currently being developed.

Table 2: Comparison of practices' self-assessment with resilience score

LMC Assessment	Practice assessment			
	Fragile	Sustainable	Thriving	
Vulnerable	12	10	2	23
Average	16	50	9	74
Resilient	1	31	7	41
	29	91	18	

3.2 Results by metric

As discussed in the previous section, 14 of the 23 metrics were found to be statistically significant and these are reported on below.

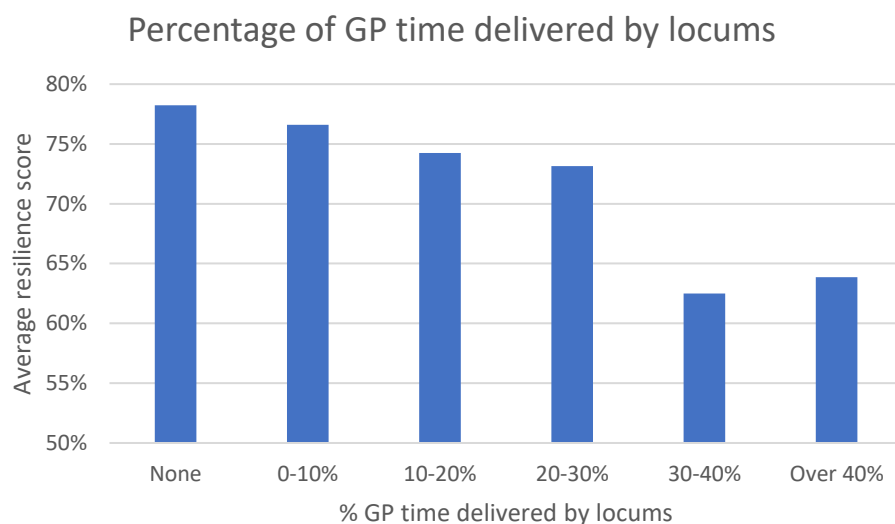
1. GPs expected to leave over the next year

Future vacancies will put additional strain on the practice and could destabilise a less resilient practice. Ninety-six practices said they were not expecting to lose any GPs and 31 were expecting to lose less than 20% of their GP hours over the next year. The average adjusted resilience score of these practices was 75% which was 13 points higher than that of the 11 practices expecting to lose 20% or more of their GP hours.

2. Locum usage

Relying on locums is less stable, more expensive and puts increased pressure on the contractors as locums tend not to carry out additional activities such as home visits or paperwork. Practices were asked how many locum sessions they had employed on average each week over the previous month. They were asked to exclude any locum sessions employed through the various winter pressure schemes that operated in some CCGs over the winter period. The number of sessions (converted into hours) was then calculated as a percentage of the overall GP weekly capacity for that practice. On average, practices were relying on locums for 15.4% of their GP sessions, but this includes 33 practices that had not used any locums over the previous month.

Practices that relied on locums for more than 30% of their GP sessions had an average adjusted resilience score that was 12 points lower than that for practices relying on locums for less than 30% (64% vs 76%). Twenty practices relied on locums for more than 30% of their GP time.

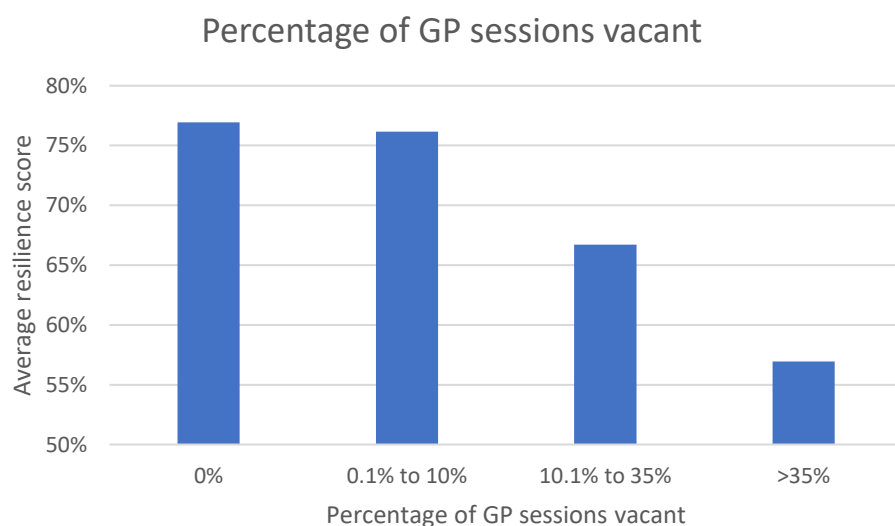


3. Current GP vacancies

Practices that are carrying a lot of GP vacancies (whether permanent or temporary due to sickness, maternity leave etc) are likely to be under greater pressure in terms of workload for the remaining staff and financial impact of employing locums.

There was a correlation between vacancy rates and overall resilience score as shown in the graph below. Eighty-two practices did not have any vacancies and 15 practices had less than 10% of their GP sessions vacant. The average adjusted resilience score for these practices was 77% which was 12 points higher than for practices where the vacancy rate was more than 10%.

Thirty-three practices had between 10% and 35% of their GP sessions vacant and their average adjusted resilience score was 67%. Eight practices had more than 35% of their GP sessions vacant and their average adjusted resilience score was 57%.





4. Home visit requests

High numbers of home visit requests puts pressure on the practice and may indicate where a practice has not been able to manage demand for home visits. The majority of practices (75 or 54%) said home visits were a problem but they managed, and 52 (38%) said home visits were not a problem. Eleven practices (8%) indicated that they struggle with home visit requests, and these practices had on average an adjusted resilience score that was 11 points lower than for the practices that did not have a problem or were managing (65% vs 76%).

5. Recruitment of GPs

Notwithstanding the GP shortage, some practices seem to be able to recruit GPs more easily than others, and some practices have not been able to replace GPs despite several attempts to recruit. Out of the 138 practices, 40 had not had to recruit GPs over the last year. The average adjusted resilience score for these practices was 78%.

Of the remaining 98 practices, 18 said they had no difficulty in recruiting GPs. The average adjusted resilience score for these practices was 84% and this was 6 points higher than for those who said they had difficulty but had managed to recruit eventually (n=37) and 15 points higher than the average adjusted resilience score for those who said they had difficulty recruiting and still had vacancies (n=43). This is shown in the graph below.

Practices that had been able to recruit had an average adjusted resilience score of 80% and this was 11 points higher than for practices that said they still had vacancies.

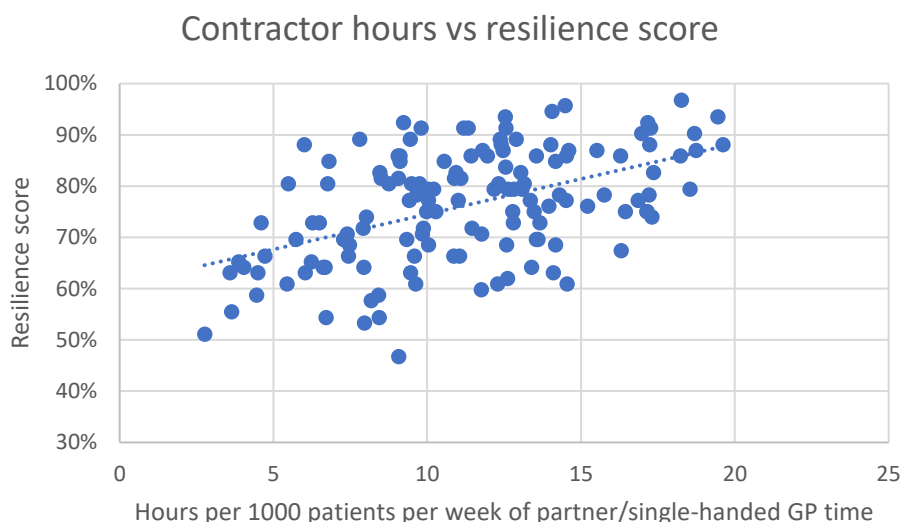


6. Contractor clinical hours in the practice

Data from the LMC Resilience Programme carried out in 2017/18 had indicated a correlation between clinical hours delivered by partners or single-handed GPs (“contractors”) and the overall resilience of the business. Data from the PHQ reinforced this correlation, and showed that practices with fewer than 9 hours per week per 1,000 patients of contractor clinical time had an average adjusted resilience score of 68% which was 11 points lower than the average adjusted resilience score of practices with 9 or more hours per 1,000 patients per week.

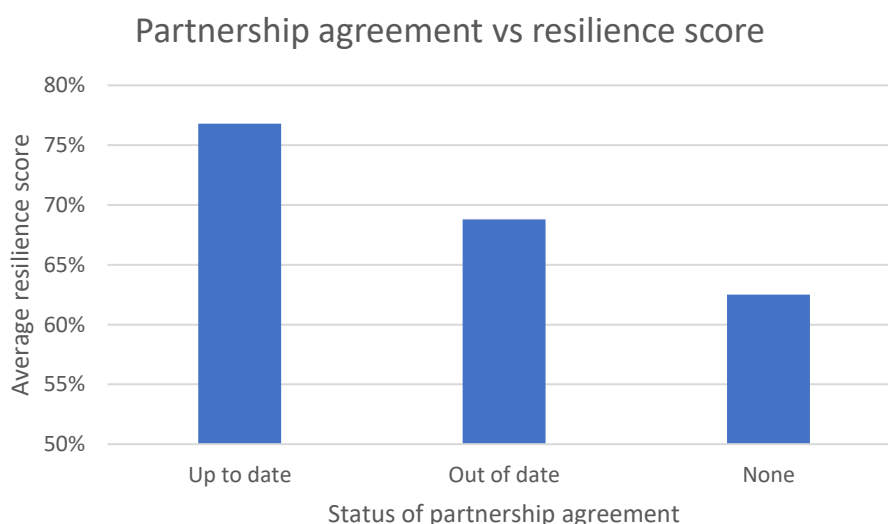
Thirty-eight (26%) of the practices had fewer than 9 hours of contractor time and an average adjusted resilience score of 68%. Fifteen (11%) practices had 17 or more hours of contractor time per 1,000

patients per week and their average resilience score was 85%. The average number of hours per 1,000 patients delivered per week across all practices was 11.1. The range is shown in the graph below.



7. Partnership agreements

Not having an up to date partnership agreement causes many problems (often expensive) when there are partnership disputes or when partners leave the practice, even where the departure is amicable. Out of the 138 practices, 124 (90%) were partnerships. Of these, 88 had an up to date partnership agreement in place, 32 had an agreement that was not up to date, and 4 said they did not have an agreement. Practices that had an up to date partnership agreement in place had an average adjusted resilience score of 77% which was 9 points higher than the average adjusted resilience score of practices with no partnership agreement or where the agreement was out of date. There was a difference of 6 points between practices with no agreement and those whose agreement was out of date but this difference was not statistically significant.

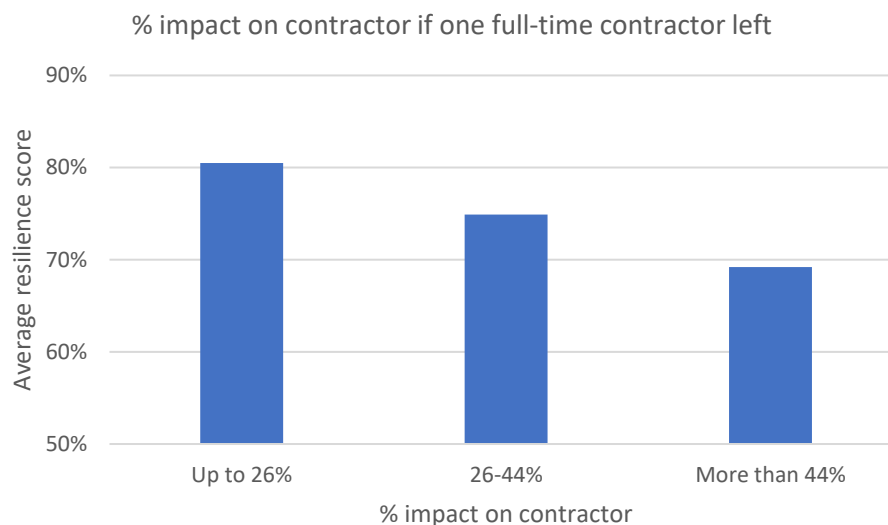




8. Impact of losing one full-time equivalent contractor

If one contractor leaves, the impact is greater the fewer partners there are. Single-handed practices are at particular risk if something happened to the sole contractor. This metric calculates the percentage reduction if one 8 session contractor leaves the practice. For single-handed practices, the impact would be 100%. The data showed that practices where the impact would be less than 26% (the equivalent of a four full-time partner practice losing one partner) had an average adjusted resilience score of 81%, and this was 8.5 points higher than where the impact would be greater than 26%.

The graph below shows difference in average adjusted resilience score between practices where the impact of losing a contractor would be less than 26%, between 26% and 44%, and above 44%.



9. List closure request

Practices can apply to the CCG to close their list on a temporary basis but they only tend to do this when under extreme pressure. The CCG does not always agree to the closure. Practices were asked whether they had applied to close their list at any time over the previous year, regardless of whether or not the application had been granted. Twenty-three practices (17%) had applied to close, and their average adjusted resilience score was 7 points lower than for the practices that had not applied to close (68% vs 75%).

10. Reduction in drawings

Reduced partner drawings can indicate a reduction in profitability and an increased risk of partners leaving or terminating their contract. However, it can also indicate that the partners have invested in their business over the previous year and reduced their drawings accordingly.

Practices that reported a drop in partner drawings over the past year (n=34, average adjusted resilience score = 70%) had, on average, a score that was 7 points below the average for practices where partners had not seen a reduction (n=104, average adjusted resilience score = 77%).

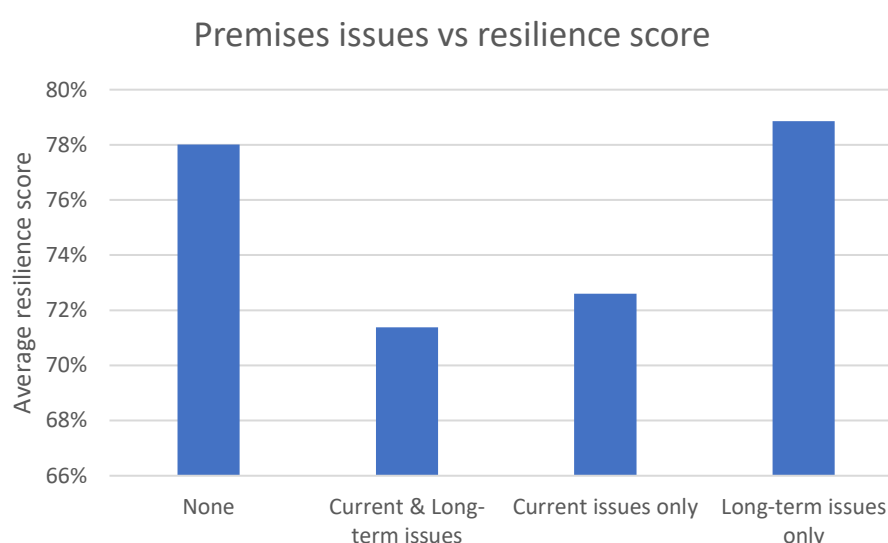
11. Premises issues

Practices were asked two questions about premises:

- whether they had any current issues that impact on the provision of services, and
- whether the long-term future of the premises was secure.

Premises issues such as capacity and quality can lead to instability and financial strain, restrict service provision, increase stress on staff and prevent new partners or staff from joining. Not being sure of the long-term future of the premises can also place strain on the partnership and prevent new partners from wanting to join. It can also prevent the CCG from agreeing to invest in premises developments or improvements.

Fifty-six (41%) practices reported that they had issues with their premises that impact on the provision of services ("current issues") and 57 (41%) had concerns about the long-term future of their premises ("long-term issues"). Thirty-seven of these had both current and long-term issues. The results showed no statistical difference in average adjusted resilience scores between practices without premises issues and those that had **either** current issues¹ **or** long-term issues (but not both). However, where a practice had both current issues **and** long-term issues, the average adjusted resilience score was 7 points lower than for practices with no premises issues and this difference was statistically significant.



12. Regular partnership meetings

This may depend on the size of the partnership, but we have seen that partnerships that meet formally on a regular basis are able to plan better and manage operational challenges. Out of the 138 practices, 124 were partnerships. Of these, 109 practices said they held formal partnership meetings at least quarterly and these had an average adjusted resilience score of 75%. Thirteen practices said they met less than quarterly or on an ad hoc basis and two said they had not met for at least a year. The average adjusted resilience score for these 15 practices was 70%, 5 points lower than for those that met quarterly.

13. Relationship with neighbouring practices

The PHQ was designed and sent out before Primary Care Networks (PCN) came in although in most areas practices were already starting to work more closely with other practices. We asked practices two questions about their relationship with other practices: one about their relationship with the practices in their immediate vicinity and one about the practices in their wider locality.

¹ The 5 point difference in score was not statistically significant



By “immediate vicinity” we meant the practices that are closest. Sometimes practices can have difficult relationships with their immediate neighbours due to historical issues, while getting on much better with practices further afield. The formation of PCNs has shown how sometimes practices co-located even in the same building do not get on with each other and have joined different PCNs.

Out of the 138 practices the majority (113 or 82%) said they had a good relationship with their neighbouring practices, and the average adjusted resilience score for these practices was 75%. This was 5 points higher than those who described the relationship as fair (22) or poor (3) and this difference was statistically significant.

14. Training practices

Training practices tend to be able to recruit from their trainees, increasing their resilience. The results showed a correlation between training practice status and resilience score. Eighty-three (60%) of the practices were training practices and had an average adjusted resilience score of 77%. The average adjusted resilience score of the remaining 55 practices was 73% and this difference of 4 points was statistically significant.

3.3 Metrics that were not statistically significant

The following metrics were found to have no statistical significance with regards to a practice’s overall resilience score:

- Having regular clinical meetings
- Having regular staff meetings
- Long term security of tenure of premises²
- Current premises issues
- Increase or decrease in number of complaints
- Relying on an overdraft to manage cash flow
- Relationships with practices in the wider locality

These metrics are not discussed further in this report.

3.4 Clinical skill mix metrics

Three further metrics were found not to be statistically significant in the way we had used them in the analysis of the PHQ but require further discussion. These are

- Number of GP and non-GP clinical specialist hours per 1,000 patients per week
- Proportion of all GP & non-GP clinical specialist hours that are delivered by GPs
- Number of nursing team hours per 1,000 patients per week

Increasingly, practices are being encouraged to employ a wider range of clinical professionals who are able to deliver some of the work that previously could only be delivered by GPs.

There were three metrics relating to clinical skill mix. Two of these looked at numbers of hours of clinician time in two groups. The first group comprised GPs and what we called “non-GP clinical specialists” which included nurses working in an advanced capacity (including minor illness nurses), pharmacists, paramedics, physicians associates and physiotherapists. The second group comprised what we called “the nursing team” which included practice nurses and healthcare assistants. We hypothesised (and scored accordingly) that the higher the number of hours of each group, the more

² As discussed above, the data showed statistically significant differences in resilience score between practices with both long-term and current premises issues.



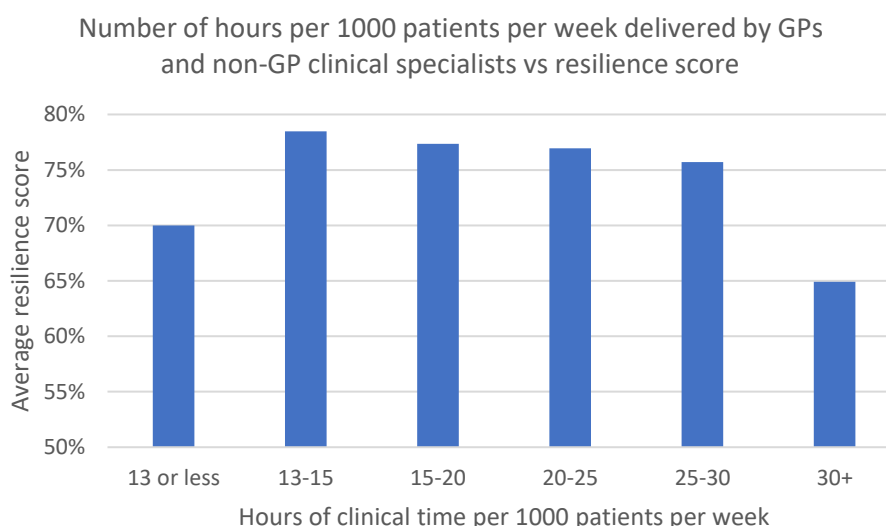
resilient the practice would be. The third metric in this category looked at the percentage of total hours of GP and non-GP clinical specialist time that was delivered by GPs. We hypothesised that more resilient practices would have a good balance, rather than being overly reliant on either GPs or non-GP clinical specialists.

Unlike the metric about contractors (i.e. Partner or Single-handed GP) hours per 1,000 patients per week, where there was a clear correlation between more hours and better resilience scores, the principle of “the more the better” did not apply in relation to any of the “clinical skill mix” metrics.

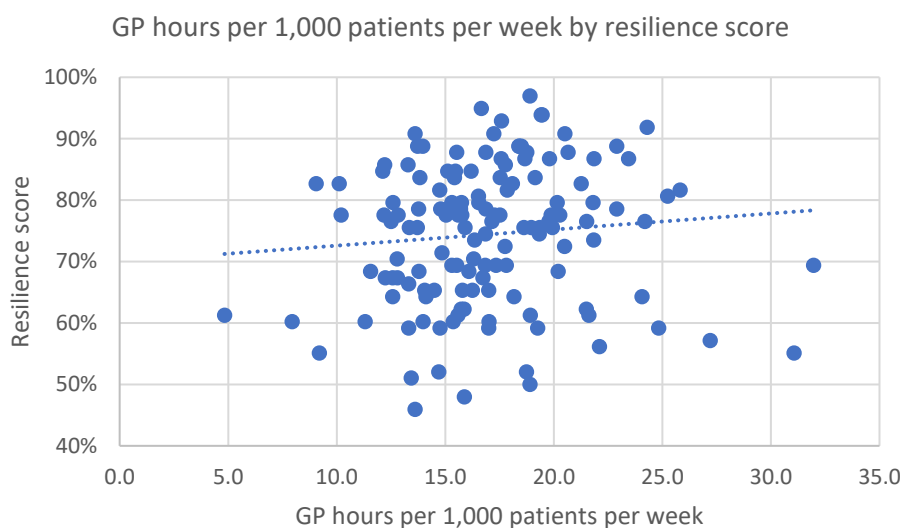
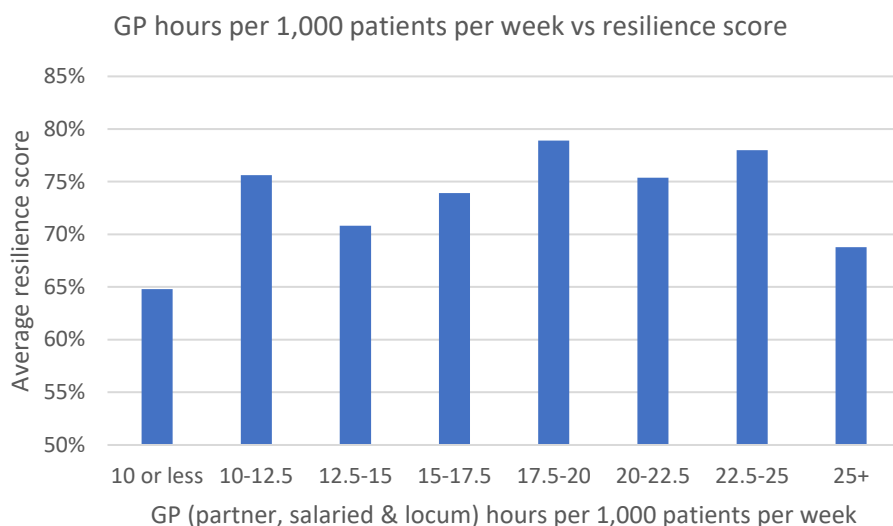
3.4.1 GPs and non-GP clinical specialists

We had grouped GPs and non-GP clinical specialists together because the view of the GP Forward View and the Long Term Plan is that these additional non-GP roles will increasingly fill roles previously filled by GPs. They are carrying out duties that are, on the whole, considered to be Essential Services within the GMS contract such as prescribing and home visits.

The metric was scored on the basis of “the more the better”. However, we found that practices that had between 13 and 30 hours per 1,000 patients per week of GP and non-GP clinical specialist time had an average adjusted resilience score of 77%. Outliers, i.e. practices with fewer than 13 hours or more than 30 hours had an average adjusted resilience score of 67% and this difference was statistically significant.



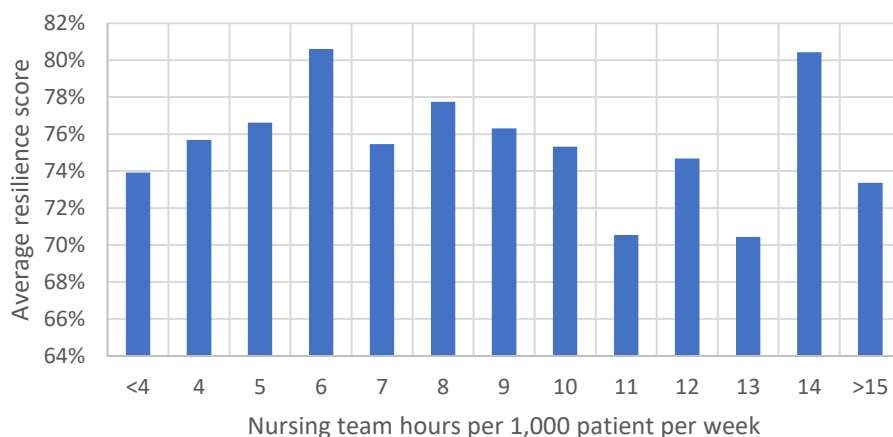
In our further analysis we separated GPs from non-GP clinical specialists to see if there were any differences. Although there was a statistically significant difference between practices that delivered more than 15 hours per 1,000 patients per week of GP time (n=97, average score = 76%) and those that delivered fewer than this (n=41, average score = 71%), when broken down further there was no correlation between number of GPs and resilience score.



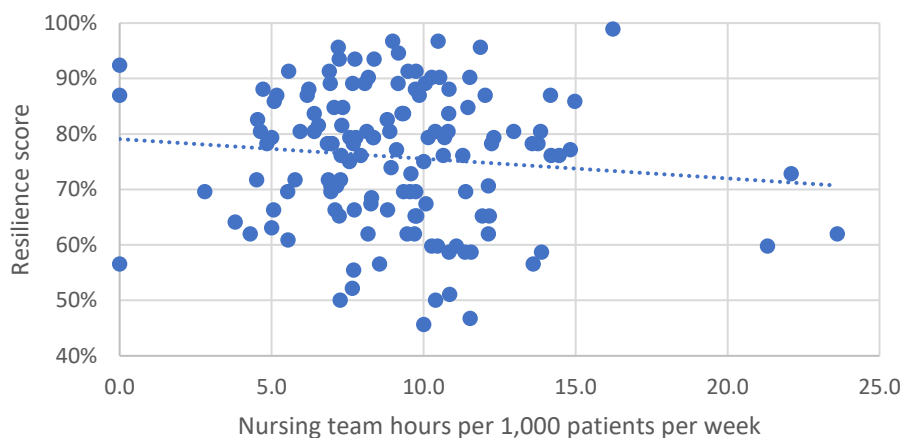
3.4.2 Nursing team

We found no difference in the average adjusted resilience scores and the number of nursing hours per 1,000 patients per week, as shown in the graph below.

Nursing team hours per 1,000 patients per week by resilience score

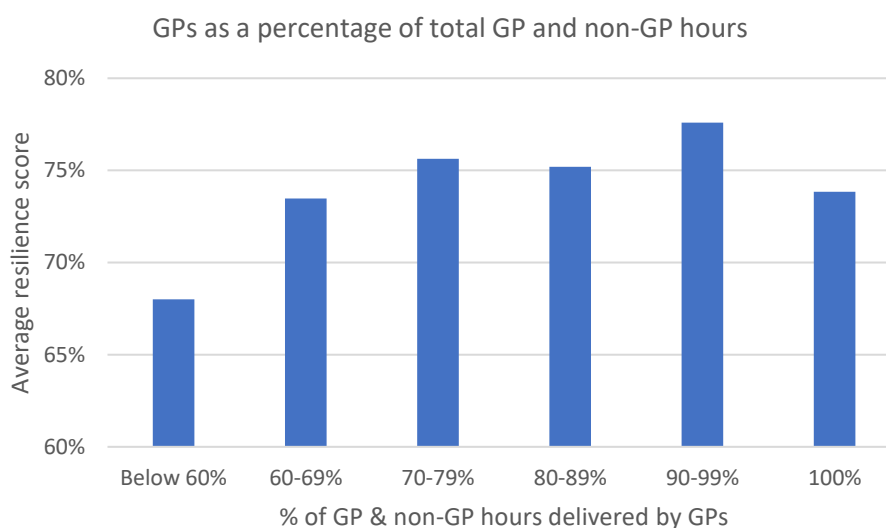


Nursing team hours per 1,000 patients per week by resilience score



3.4.3 Percentage of total GP & non-GP hours delivered by GPs

Because there were no benchmarks for clinical skill mix, we originally set the scoring for this metric based on the average. The average percentage of total GP & non-GP hours that were delivered by GPs was 82% and originally we set the scoring so practices that were around the average scored highest, and outliers above or below this scored lower. However, we were uncomfortable with this scoring mechanism as it was not based on any evidence and so decided to remove this metric from the analysis altogether. As part of our further analysis, we were able to show that there is a correlation between practices where non-GPs deliver a high percentage of the total GP & non-GP hours but we are not able to show whether or not this is statistically significant.



3.5 New metrics identified

In testing the metrics for statistical significance, we identified some additional metrics that we had not expected. We will consider how to incorporate these findings in future iterations of the PHQ.

Training practices without trainees

As well as identifying a statistically significant difference between training practices and non-training practices, we also identified a difference between training practices that currently have trainees, and training practices that do not have trainees. 83 (60%) of the practices were training practices, and 5 of these said they currently did not have any trainees. The average resilience score of the training practices with trainees was 78% and of those without was 63%, a difference of 15 points and this difference was statistically significant.

Difficulty in recruiting doctors

Practices were given four answer options to the question “Over the last year, how difficult has it been to recruit GPs?”. In terms of scoring, we gave practices that had not had to recruit the same score as practices that had not had any difficulty in recruiting. However, on further analysis of the data, we found that practices that had been able to recruit GPs without difficulty scored, on average, 6 points higher than those that had not had to recruit (84% vs 78%), and this difference was statistically significant.

Practice self-assessment

We asked practices whether they considered themselves to be thriving, sustainable or fragile. Their response to this question was not included in the formation of the metrics. However further analysis showed that practices that assessed themselves as thriving or sustainable³ (n=109) had an average score of 77% which was 13 points higher than those that assessed themselves as fragile (n=29) who

³ There was no significant difference in average score between practices assessing themselves as thriving and those assessing themselves as sustainable.



had an average score of 64%, suggesting that a practice's self-assessment should be taken into account in future particularly where they assess themselves as being fragile. Having said that, it should be noted that one practice that scored above 90% assessed itself as fragile.

List size

We found a correlation between list size and overall resilience score as shown in the graph below. The resilience score increased with list size, but this was not entirely incremental.



We found that the difference in average score between practices with a list size of below 9,000 (n=48, average score = 70.3%) was 6 points lower than the average score for practices with a list size above 9,000 (n=90, average score = 76.6%) and this difference was statistically significant.

Seven practices had a list size above 21,000. Their average resilience score was 81% but although this was 5 points higher than the average score of the practices with a list size of between 9,000 and 21,000 (n=83), this difference was not statistically significant.



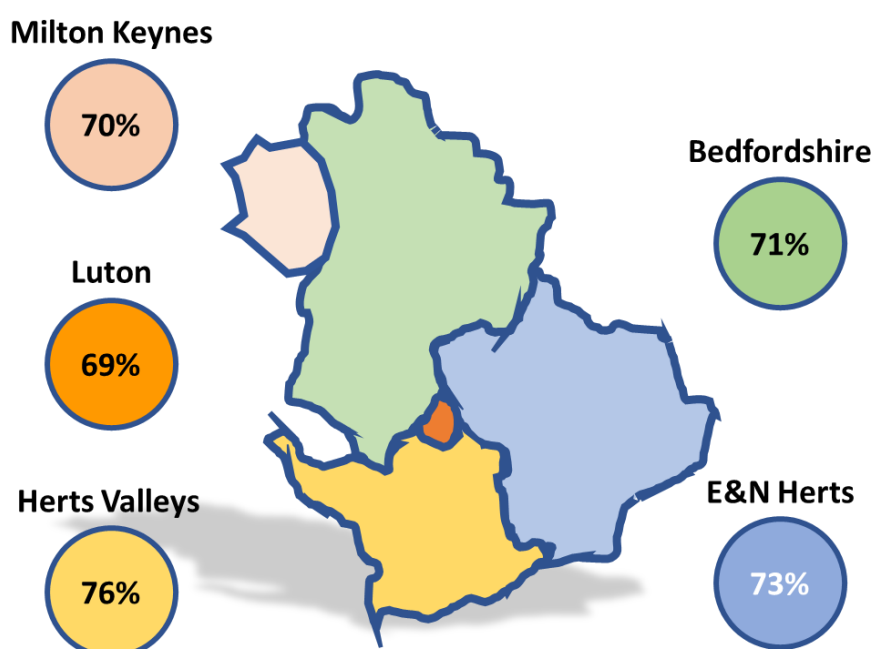
Section 4: Results by CCG

The PHQ was sent to all GMS practices in Bedfordshire and Luton in September 2018 and to all GMS practices in Hertfordshire and Milton Keynes in January 2019. The table below shows the number of GMS practices in each CCG that completed the PHQ and the overall average resilience score. The figure overleaf shows the average resilience scores for each CCG

Table 1: Responses by CCG

	Total GMS	Returned PHQ	% return
Bedfordshire	43	30	70%
Luton	23	21	91%
Milton Keynes	23	17	74%
East & North Herts	49	36	72%
Herts Valleys	56	34	60%
TOTAL	196	138	70%

Figure 2: Average Resilience Scores for practices by CCG



The differences in resilience score between the practices in CCGs within BLMK STP and those in the Hertfordshire part of HWE STP are not surprising considering the different situation in general practice in those areas. Since 2013, Bedfordshire and Luton CCGs have had 12 contracts handed back or terminated (14% of the 86 practices that were in place at the end of 2013), mainly as a result of partners leaving or retiring. This is compared with 8 in Herts Valleys and East and North Herts CCGs (6% of the 130 practices that were in place at the end of 2013). In Bedfordshire and Luton, 7 of the 12 were closed and the patients dispersed, 4 are now APMS and 1 is still being sorted out. In Herts Valleys and East and North Herts, 2 of the 8 were closed and dispersed (both of these had been APMS practices before being closed), and 6 are now run as APMS. Please note that the changes described in this paragraph do not take into account practice mergers which have also taken place in both STP areas over the last few years.

4.1 CCG results by metric

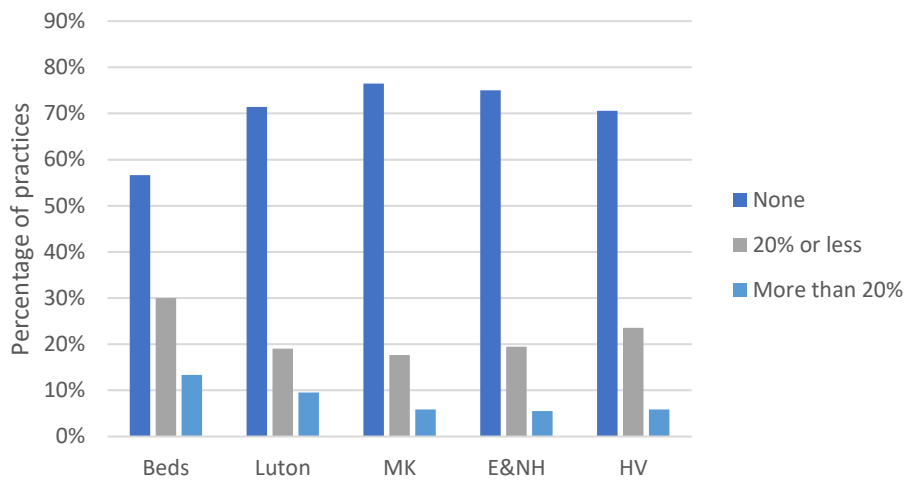
The PHQ used 23 metrics to assess the resilience of a practice of which 14 were found to have a statistically significant impact on overall resilience. These 14 are reported on below.

1. GPs expected to leave over the next year

Practices that were expecting to lose more than 20% of their GP workforce over the coming year had on average resilience scores that were lower than those expecting to lose 20% or less. Forty-two practices were anticipating losses over the next year. This is shown by CCG in the graph below.



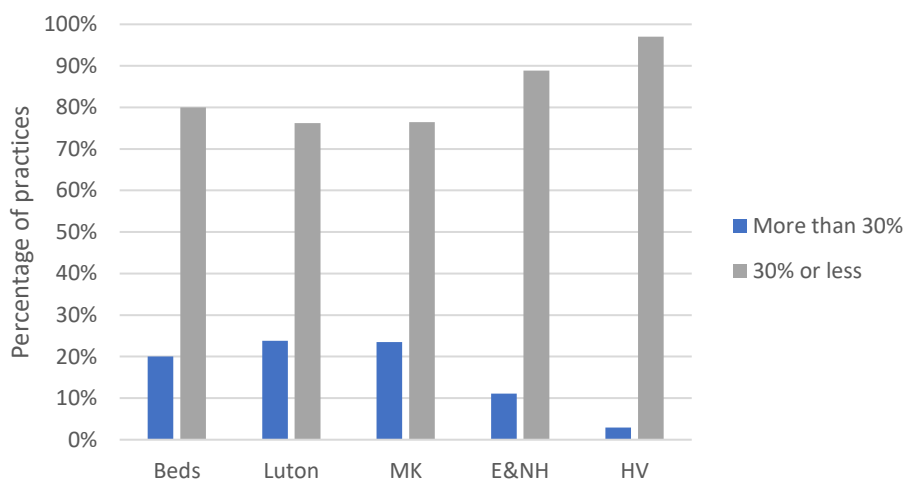
GPs expected to leave over the next year



2. Locum usage

The data showed that practices that relied on locums for 30% or less of their GP sessions had higher resilience scores than practices that relied on locums for more than 30%. Locum usage was higher in the BLMK CCGs than in Hertfordshire.

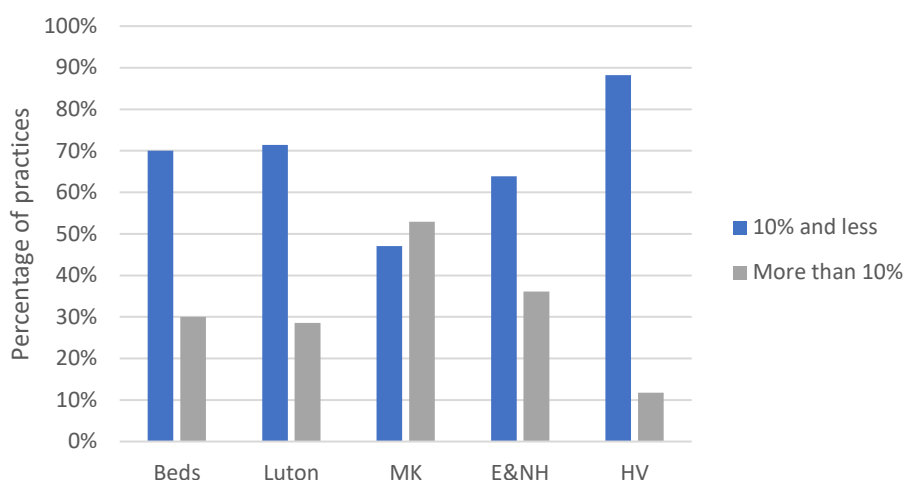
Locum use as a % of total GP provision



3. Current GP vacancies

The data showed that practices carrying more than 10% of vacancies had on average a resilience score that was lower than practices with fewer than 10%. This information is compared by CCG in the graph below.

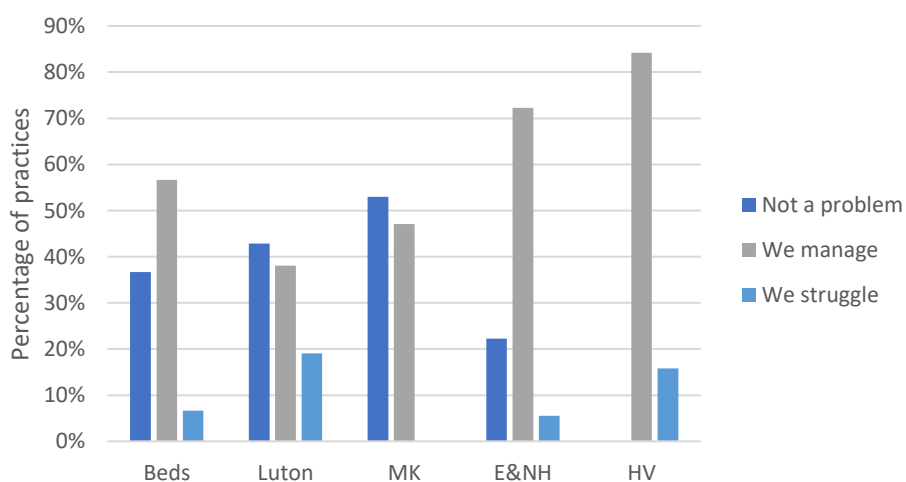
GP vacancies



4. Home visit requests

The data showed that practices that say they struggle with home visits had an average resilience score that was lower than those who said they didn't have a problem with home visits or that they managed. This is shown by CCG in the graph below.

Difficulty with home visits

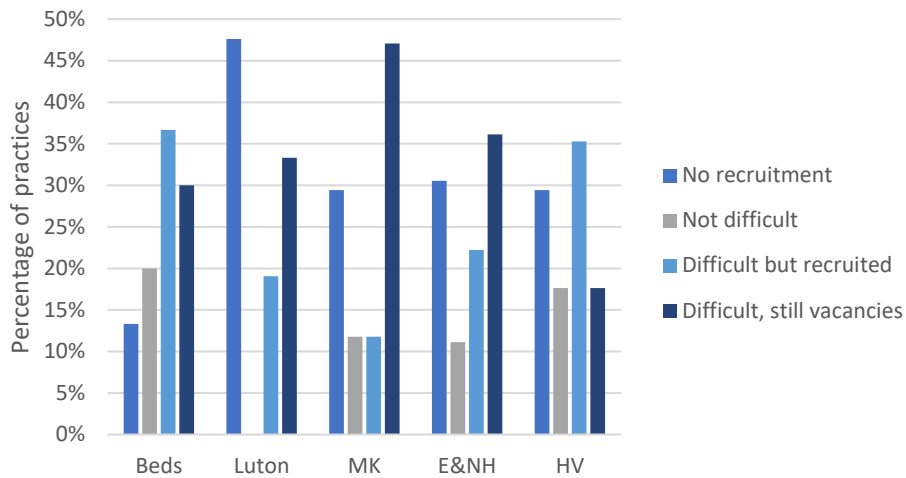


5. Being able to recruit GPs

Some practices found it easier to recruit new GPs than others. Practices that found it difficult and still have vacancies had on average lower resilience scores than those who had been able to recruit. This is shown by CCG in the graph below. This shows that this is particularly an issue in Milton Keynes.

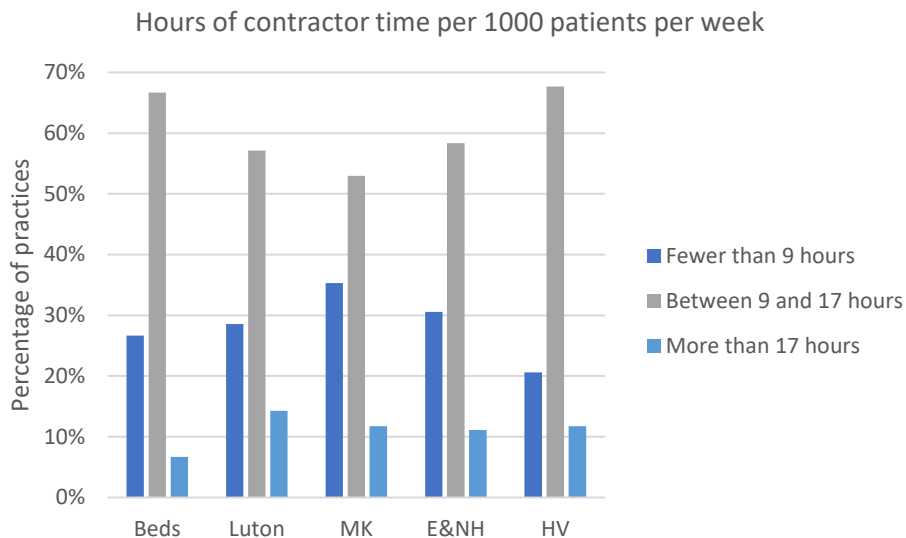


Difficulty in recruiting GPs over the last year



6. Contractor clinical hours in the practice

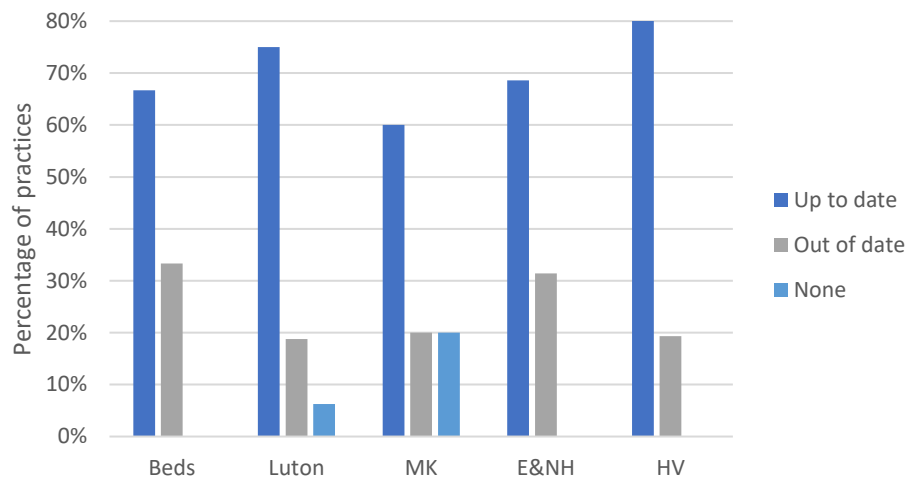
Practices where contractors (partners or single-handed GP) deliver fewer than 9 hours per 1,000 patients per week had resilience scores that were significantly below those for practices delivering more than 9 hours a week. This is shown by CCG below.



7. Partnership agreements

The data showed a correlation between having an up to date partnership agreement and having a higher resilience score. This is shown by CCG in the graph below.

Status of partnership agreement

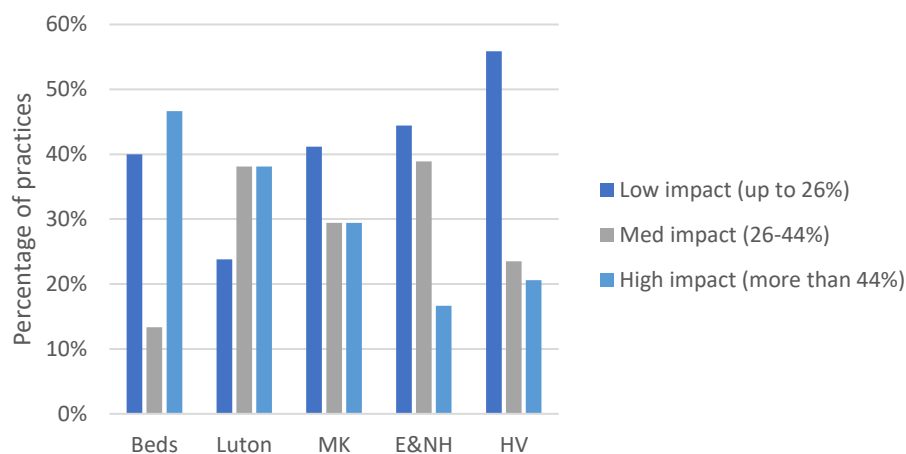


8. Impact of losing one full-time equivalent contract holder

The data showed a correlation between resilience and the percentage impact on the partnership if one full-time equivalent partner left. For single-handed GPs, the impact would be 100%. Practices where the impact would be less than 26% had on average a resilience score that was statistically significantly higher than those where the impact would be greater than 26%.

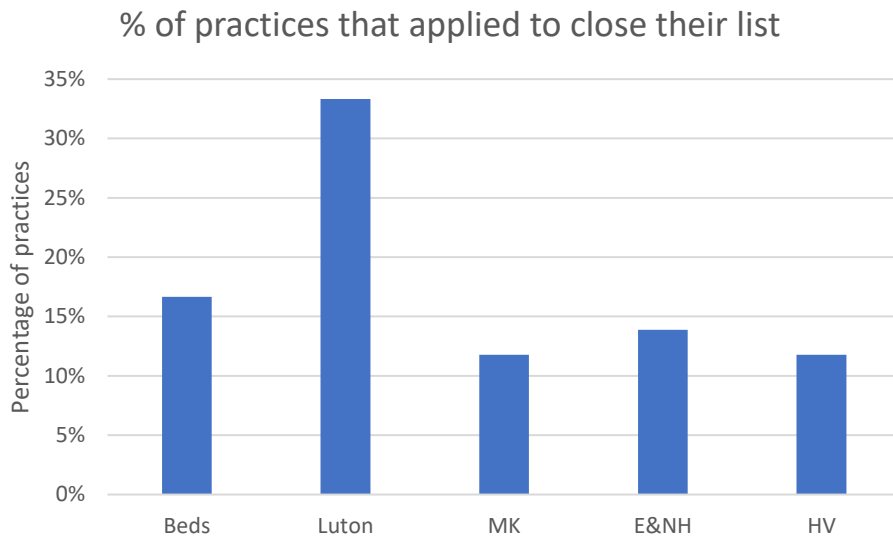
24% of the practices in Luton that completed the PHQ are single-handed compared with 12% in Milton Keynes, 10% in Bedfordshire, 9% in Herts Valleys and 3% in East & North Herts.

Impact of losing a full-time equivalent contractor as a % of total contractor hours per week



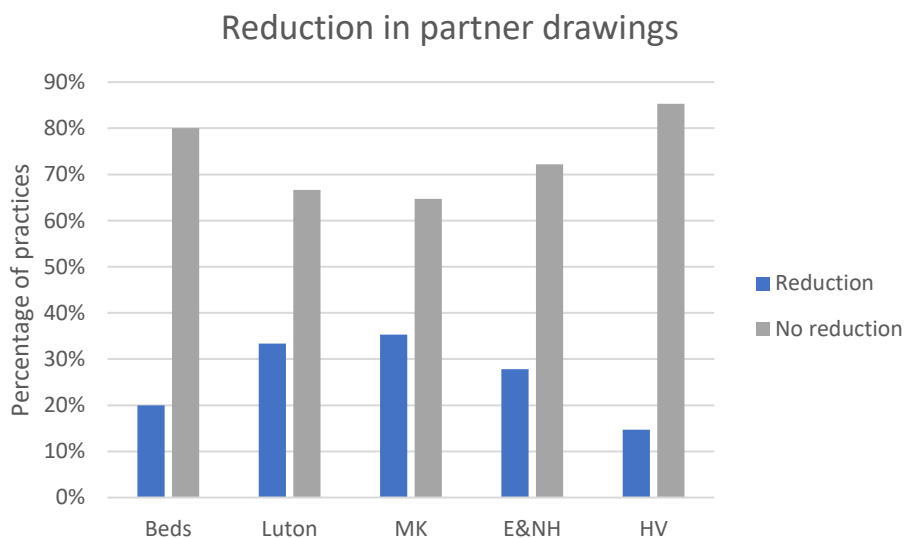
9. List closure request

In total 23 practices had applied to close their list over the last year. Seven of these were in Luton (30% of the Luton practices that completed the PHQ).



10. Reduction in drawings

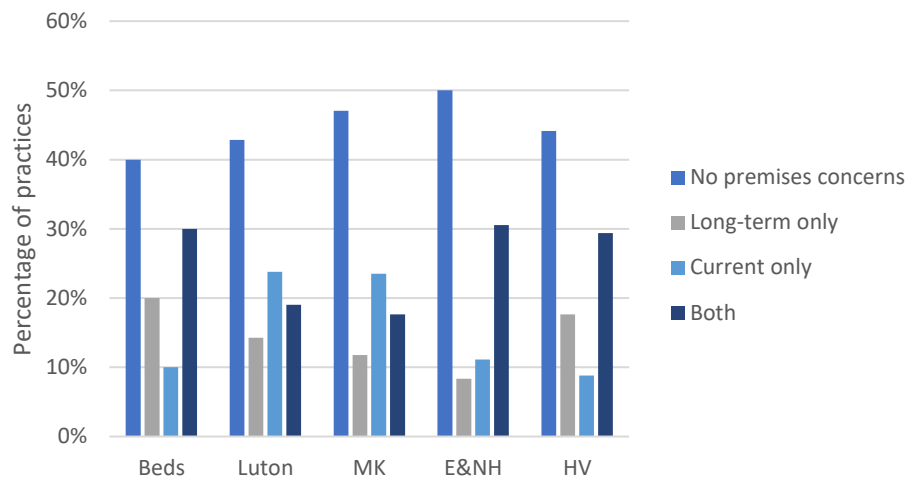
The data showed that practices where the partners' drawings had reduced had on average lower resilience scores than for practices where the drawings had not gone down. The graph below shows the percentage of practices in each CCG that reported a reduction in partner drawings.



11. Premises issues

The data showed that practices with both current premises issues and concerns about the long-term future of their premises had on average lower resilience scores than those with no premises issues. These are indicated in dark blue in the graph below which shows premises issues by CCG.

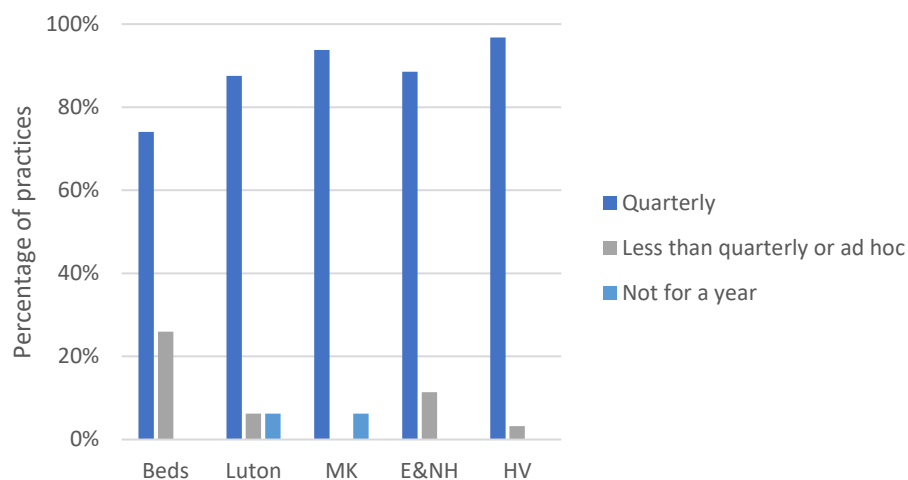
Issues related to premises



12. Regular partnership meetings

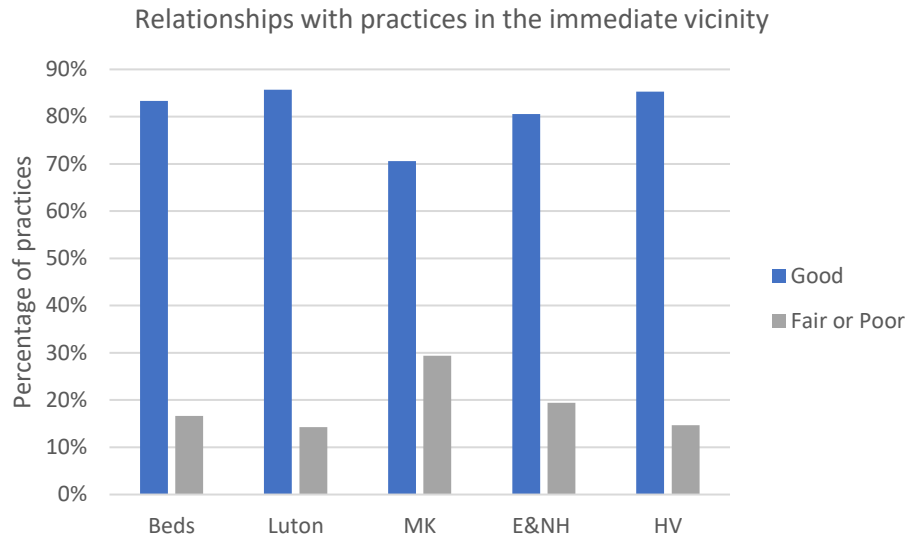
The majority of partnerships said they meet formally at least quarterly. Partnerships that met less than quarterly or on an ad hoc basis (or, in two cases, had not met for a year), had on average lower resilience scores than where the partners met regularly. This is shown by CCG below.

Frequency of formal partnership meetings



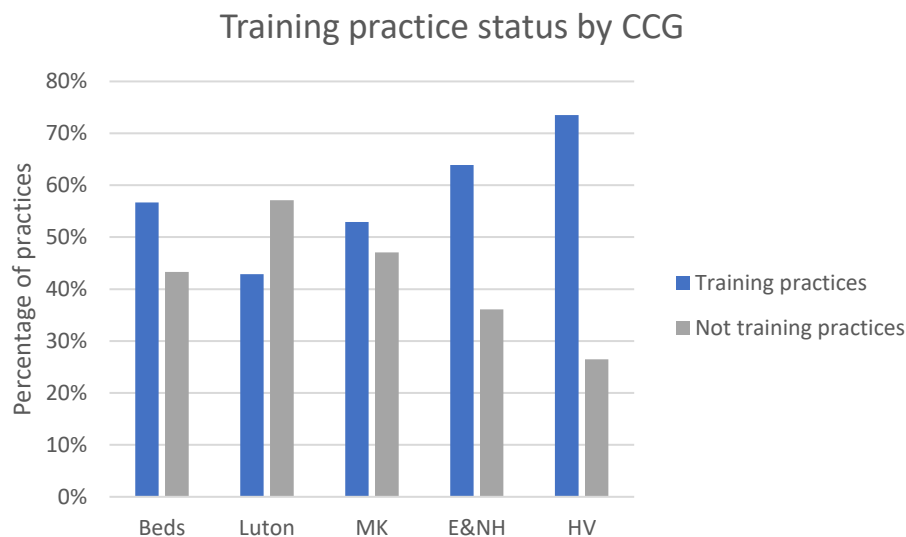
13. Relationship with neighbouring practices

The majority of practices reported they had good relationships with their immediate neighbouring practices, though 25 said the relationship was fair or poor, and these practices had on average lower resilience scores than where the relationships were good. This is shown by CCG in the graph below.



14. Training practices

Training practices had, on average, slightly higher resilience scores than non-training practices. The graph below shows training practice status for the practices that completed the RAT by CCG. The data showed that training practices that currently do not have any trainees had statistically significant lower resilience scores on average than both training practices with trainees and non-training practices. There was 1 training practice in Bedfordshire and 2 each in East & North Herts and Herts Valleys that do not currently have any trainees. The scores for these five practices ranged from 52% to 69%.



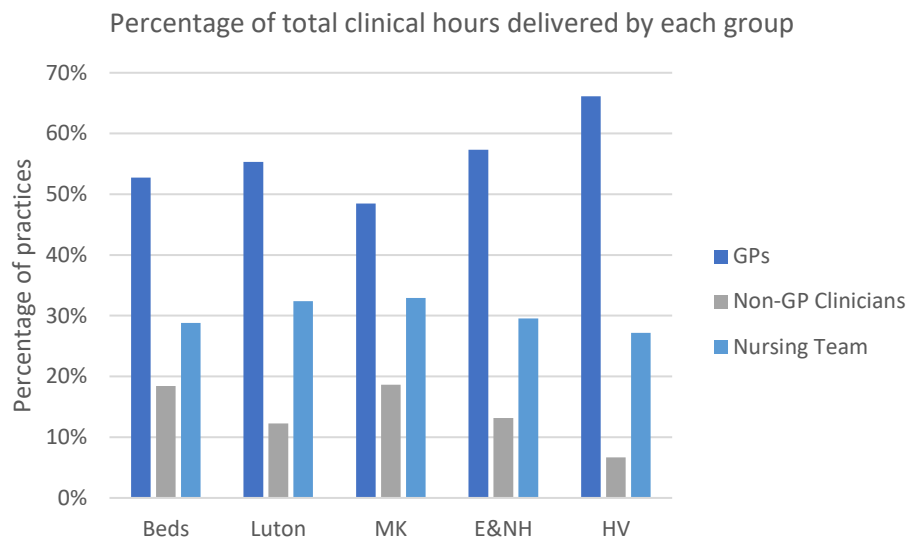
15. Clinical Skill Mix

The data did not show a statistically significant relationship between clinical skill mix and resilience score. However, there is a great emphasis in the GP Forward View and in the NHS Long Term Plan on diversifying the workforce in primary care and employing more non-GP clinical specialists to carry out



some of the work previously done by GPs as a way to address the shortage of GPs. Therefore, this information is of interest to CCGs.

The graph below compares average skill mix between CCGs. This shows a greater use of non-GP clinical specialists in Bedfordshire and Milton Keynes, perhaps as a reflection of the difficulties in recruiting GPs in those areas, or as a result of greater promotion of non-GP roles in those CCGs. As reported above, however, the skill mix within a practice had no statistically significant impact on overall resilience.





Section 5: Next Steps

Having completed the analysis of the data from 138 practices, we are now planning the next stage of this project.

Review of the PHQ

Taking into account the findings from the data analysis, we intend to make some changes to the PHQ in terms of the questions asked, the metrics used and the scoring system. In particular, the changes we intend to make include:

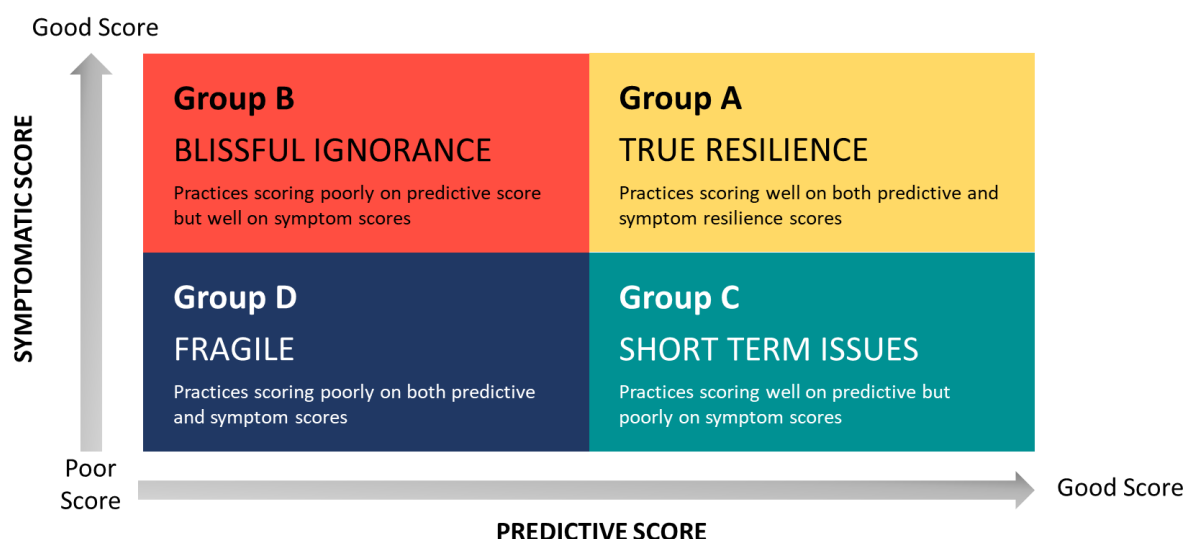
- Removing questions that had no relevance to the resilience of a practice, such as questions about complaints or staff meetings.
- Changing the weighting to reflect the relative impact, for example reduce the weighting attributed to training practices.
- Adding new questions such as with regards to practice management
- Considering whether to create new metrics based on the findings outlined in Section 3
- Considering whether continue to include metrics about clinical skill mix and how these may need to be amended
- Adding a facility to be able to analyse and report the data by PCN.

We would then like to invite practices to complete the revised questionnaire.

Further analysis of the data

Some of the metrics we used seemed to reflect the infrastructure of the practice (e.g. being a training practice, having an up to date partnership agreement) and could be used to predict whether or not a

practice is resilient. Other metrics could be considered symptoms of fragility, i.e. a sign that the practice is under pressure, such as relying heavily on locums. We started to look at dividing the metrics into those that are “predictive” and those that are “symptomatic”, and categorised practices into one of four groups as shown in the diagram below. In this way, we thought we could identify practices needing support in a more nuanced way than merely looking at their overall resilience score.



However, it became clear that some metrics could be both the cause of a practice’s relative resilience and the effect. For example, having fewer than 9 hours per 1,000 patients per week of contractor time could be a “predictive” metric if it was a conscious decision by the partners not to increase the partnership, or it could be a “symptomatic” metric where partners have left and the practice has been unable to replace them.

We would like to explore these ideas further before being able to make any firm conclusions about whether or not we could use the metrics to reliably identify a practice’s specific situation with regard to their predictive and symptomatic resilience. However, we believe that this method of categorisation of a practice’s resilience may be useful in identifying the most appropriate support package for them. Currently, support offered by CCGs, STPs, and LMCs frequently follows a “one size fits all” model, often focusing on providing GP back-fill to support struggling practices. While this may be appropriate for some practices for a short period of time (specifically those practices in Group C), practices in the other quadrants may benefit more from support directed at addressing some of the underlying issues effecting their resilience as a business.

List of metrics in order of statistical significance (most significant first)

Metric	Rationale/Assumption	Original weighting	Outcome
% of GP posts currently vacant	<i>Carrying a lot of GP vacancies puts strain on the remaining GPs and financial strain on the business.</i>	High	On average, practices where 10% or less of the GP posts were currently vacant had a resilience score that was 12 points higher than practices where the vacancy rate was above 10% (77% vs 65%). This difference was statistically significant.
Up-to-date partnership agreement in place	<i>Not having an up-to-date partnership agreement causes many problems (often expensive) when there are partnership disputes or partners leave the practice.</i>	High	<p>On average, practices that had an up to date partnership agreement had a resilience score that was 9 points higher than practices whose partnership agreement was out of date or that did not have a partnership agreement at all (77% vs 68%). This difference was statistically significant.</p> <p>There was no statistical difference in score between practices that had an out-of-date agreement and those who did not have an agreement at all (70% vs 62.5% but the difference was not statistically significant).</p>
Relationship with practices in their immediate vicinity	<i>By "immediate vicinity" we are talking sub-locality/cluster, the practices that are closest. Sometimes practices get on very well with practices further away but badly with their immediate neighbours and this can cause problems. This metric takes on more significance now that practices are working in networks with other practices.</i>	Medium	<p>On average practices that had a good relationship with the practices in their immediate vicinity had a resilience score that was 14 points higher than where the relationship was poor (75% vs 61%). Although this difference was statistically significant, only 3 practices out of 138 stated that the relationship with their neighbours was poor.</p> <p>The difference in average score between practices with a good relationship and those who described the relationship as poor or fair (n=25) was 5 (75% vs 70%), and this difference was also statistically significant.</p>
% of GP sessions delivered by locums	<i>Relying on locums is less stable and puts increased pressure on the partners/contract holder as locums tend not to carry out additional activities such as home visits or paperwork as well as being expensive.</i>	High	On average, practices where locums delivered less than 30% of the GP sessions had a resilience score that was 12 points higher than where locums delivered 30% or more (76% vs 64%). This difference was statistically significant.
% of GPs expecting to leave over the next year	<i>Future GP vacancies will put additional strain on the practice and could destabilise a less resilient practice.</i>	Medium	On average, practices that were expecting to lose less than 20% of their GP hours over the next year had a resilience score that was

Metric	Rationale/Assumption	Original weighting	Outcome
			13 points higher than practices expecting to lose 20% or more (75% vs 62%). This difference was statistically significant.
Number of clinical hours delivered by partner/single hander per 1,000 pts per week	<i>The RAT data indicated a correlation between clinical hours by partners/contract holder into the business and overall resilience of the business</i>	High	On average, the resilience score for practices with more than 9 hours of clinical time delivered by the partners/contract holder per 1000 patients per week was 11 points higher than those with fewer than 9 hours per week (79% vs 68%). This difference was statistically significant.
Regular partnership meetings	<i>This may depend on the size of the partnership, but we have seen that partnerships that meet formally are able to plan better</i>	High	<p>On average, practices that had held formal partnership meetings at least quarterly had a resilience score that was 12 points higher than practices that had not met for at least a year (75% vs 63%). This difference was statistically significant but only 2 practices said they had not met for at least a year.</p> <p>The difference in average resilience score between practices where the partners met at least quarterly and where they met on an ad hoc basis or hadn't met for at least a year was 5 points (75% vs 70%) and this difference was statistically significant.</p>
The impact of losing one full-time contract holder (i.e. partner or single hander)	<i>If one partner leaves, the impact is greater the fewer partners there are; single-handed practices are at risk if something happened to the contract holder.</i>	High	The average resilience score of practices where the loss of a wte partner/contract holder would result in a 44% or greater reduction in the total number of partners/contract holders was 9 points lower than in practices where the reduction would be less than 44% (69% vs 78%). This difference was statistically significant.
Difficulty in recruiting doctors	<i>Notwithstanding the GP shortage, some practices seem to be able to recruit GPs while others struggle. The assumption was that more resilient practices would find it easier to recruit.</i>	Medium	<p>40 practices had not had to recruit GPs over the last year. Their average score was 78%.</p> <p>Of the remaining 98, 18 had no difficulty recruiting GPs. Their average resilience score was 84% and this was 15 points higher than the average score of those who said they had had difficulty recruiting and still had vacancies (n=40) and 6 points higher than the score of those who had had difficulty but managed to recruit eventually (n=37).</p>

Metric	Rationale/Assumption	Original weighting	Outcome
Difficulty in delivering home visits (inc care homes)	<i>High number of home visits puts pressure on the practice, and may indicate where a practice has not been able to manage demand for home visits</i>	Medium	On average, the resilience score for practices that said they struggled with home visit requests was 11 points lower than those that said home visits were not a problem or were a problem but they managed. This difference was statistically significant.
Recent list closure (or application to close list)	<i>Practices only apply to close their lists when under extreme pressure</i>	High	On average, the resilience score for practices that had not applied to close their list over the last year was 7 points higher than those that had (whether or not their application had been granted). This difference was statistically significant.
Relationship with practices in their locality/cluster	<i>This is more about how the practice works within a larger grouping, and/or within a federation. If a practice has a poor relationship with the other practices in their locality/cluster, this could hinder them from developing particularly if they are small.</i>	Medium	On average practices that had a good relationship with the practices in their locality had a resilience score that was 7 points higher than where the relationship was poor. This difference was statistically significant.
Reduction in drawings	<i>Reduced partner drawings can indicate a reduction in profitability and an increased risk of partners leaving or terminating the contract.</i>	Medium	On average, practices where partner drawings had reduced over the last year had a resilience score that was 7 points lower than where drawings had not reduced (70% vs 77%). This difference was statistically significant.
Being a training practice	<i>Training practices tend to be able to recruit from their trainees, increasing their resilience</i>	Medium	On average, the resilience score for training practices was 4 points higher than non-training practices and this was statistically significant (77% vs 73%). We also found that the average resilience score for training practices that currently do not have any trainees was 15 points lower than training practices that do have trainees (63% vs 78%), and this difference was statistically significant.
Change in the number of complaints	<i>Increasing number of complaints can indicate a practice under pressure, and having to deal with increasing number of complaints increases pressure on the practice</i>	Low	Unexpectedly, on average, the resilience score for practices whose complaints had gone down was 3 points lower than those whose complaints had gone up, but this difference was not statistically significant.
Nursing team (practice nurses and HCAs) hours per 1,000 pts per week	<i>This is to benchmark practices – assumption is that the more nursing hours per 1,000 patients the better</i>	Medium	There was no significant difference in the average resilience scores between practices offering higher or lower than average number of nursing team hours per 1000 patients per week.
Total GP & clinical specialist (e.g. ANP, pharmacist etc.) hours per 1,000 pts per week	<i>This is to benchmark practices – assumption is that the more clinical hours per 1,000 patients the better</i>	Medium	There was no significant difference in the average resilience scores between practices offering higher or lower than average number of GP and clinical team hours per 1000 patients per week.

Metric	Rationale/Assumption	Original weighting	Outcome
% of total GP & clinical specialist hours delivered by GPs	<i>This is to benchmark practices against how much they are diversifying the workforce. Assumption is that a balance between GPs and non-GPs would be best.</i>	Medium	There was no significant difference in the average resilience scores between practices that offered more GP time in relation to other non-GP clinical specialists and those offering less.
Clinicians meet informally regularly	<i>This is based on the assumption that it is good for morale and support for clinicians within the practice to have opportunities to meet regularly on an informal basis.</i>	Medium	There was no statistically significant difference in resilience score between practices where clinicians met informally and those where they did not. N.B. This question specifically asked about clinicians being able to meet informally and not about formal clinical meetings. However, it's possible that this may have been interpreted differently by those answering the questionnaire and will be reworded in future.
Regular staff meetings	<i>This is based on the assumption that a practice where staff are kept informed and have a chance to meet will function better</i>	High	There was no statistically significant difference in resilience score between practices that held regular staff meetings and those that did not.
Regular use of an overdraft (or debt) to manage their cashflow	<i>Relying on a regular overdraft or debt is not a sound business model and can lead to instability</i>	Low	There was no statistically significant difference in resilience score between practices that relied on an overdraft to manage cash flow and those that did not.
Premises issues	<i>Premises issues (such as capacity and quality) can impede the ability to provide services and lead to instability and financial strain, and prevent new partners or staff wanting to join</i>	High	There was no statistically significant difference in resilience score between practices that had current premises issues that limited their capacity to provide services, and those that did not. However where practices had both current premises issues and long-term issues the average resilience score was 7 points lower than for practices with no issues.
Long-term security of tenure	<i>Lack of long-term security of tenure can make it difficult to plan for the future, and to recruit new partners</i>	Low	There was no statistically significant difference in resilience score between practices where the long-term future of the premises was secure and those where it was uncertain. However where practices had both current premises issues and long-term issues the average resilience score was 7 points lower than for practices with no issues.

