Deloitte.

PLAN. PREPARE. PERFORM.

Demand and Capacity Review of South East Coast Ambulance Service NHS Foundation Trust

Final Summary Report

August 2018

Important notice

Deloitte LLP ("Deloitte") and Operational Research in Health ("ORH") have prepared the final summary report ("Final Summary Report") from the final report (the "Final Report") for NHS North West Surrey Clinical Commissioning Group, NHS Swale Clinical Commissioning Group, NHS Horsham and Mid Sussex Clinical Commissioning Group and South East Coast Ambulance Service NHS Foundation Trust ("SECAmb") (together, the "Client") on the terms set out in the contract between Deloitte and the Client dated 23 October 2017 ("the Contract") and the Change Orders dated 21 February 2018 and 28 June 2018 and on the basis of the scope and limitations set out below. The Final Summary Report should be read in conjunction with the Final Report, which has been provided separately to the Client and is available upon request. References to the Final Report in this Important Notice apply also to this Final Summary Report.

The Final Report has been prepared solely for the purposes of informing the contracting discussions between SECAmb and its Commissioners, as set out in the Contract. It should not be used for any other purpose or in any other context, and Deloitte accepts no responsibility for its use in either regard.

The Final Report is provided exclusively for use under the terms of the Contract by the Client and those organisations that have agreed to enter the Contract through beneficiary access agreements (the "Beneficiaries"). No party other than the Client and Beneficiaries, including the regulator, is entitled to rely on the Final Report for any purpose whatsoever and Deloitte accepts no responsibility or liability or duty of care to any party other than Client and Beneficiaries in respect of the Final Report or any of its contents. If any party other than the Client or Beneficiaries choose to rely on the Final Report, they do so at their own risk and without recourse to Deloitte.

The information contained in the Final Report has been compiled by Deloitte and Operational Research in Health and includes material obtained from information provided by SECAmb, discussions with management of the Trust, Commissioners, and Regulators, and published sources which Deloitte use regularly but has not been verified. This document also contains confidential material proprietary to Deloitte.

As set out in the Contract, the scope of our work has been limited by the time, information and explanations made available to us. The information contained in the Final Report has been obtained from the Clients and third party sources that are clearly referenced in the appropriate sections of the Final Report. Deloitte has neither sought to corroborate this information nor to review its overall reasonableness. Further, any results from the analysis contained in the Final Report are reliant on the information available at the time of writing the Final Report and should not be relied upon in subsequent periods.

This document also includes certain statements, estimates and projections provided by the Client with respect to anticipated future performance. Such statements, estimates and projections reflect various assumptions concerning anticipated results and are subject to significant business, economic and competitive uncertainties and contingencies, many of which are or may be beyond the control of the Client. Accordingly, there can be no assurance that such statements, estimates and projections will be realised. The actual results may vary from those projected, and those variations may be material. No representations are or will be made by any party as to the accuracy or completeness of such statements, estimates and projections or that any projection will be achieved.

Any decision to invest, conduct business, enter or exit the markets considered in the Final Report should be made solely on independent advice and no information in the Final Report should be relied upon in any way by any third party. This Final Report and its contents do not constitute financial or other professional advice, and specific advice should be sought about your specific circumstances. In particular, the Final Report does not constitute a recommendation or endorsement by Deloitte to invest or participate in, exit, or otherwise use any of the markets or companies referred to in it. To the fullest extent possible, Deloitte and the Client and Beneficiaries disclaim any liability arising out of the use (or non-use) of the Final Report and its contents, including any action or decision taken as a result of such use (or non-use).

Executive summary

Executive summary (1/8)

The Demand and Capacity Review estimates additional resources required to meet new national performance standards for 999 services

Background

South East Coast Ambulance Service NHS Foundation Trust (SECAmb, also the "Trust" / "Provider") faces operational and financial challenges in line with other ambulance Trusts in England.

In the context of operational and financial challenges, SECAmb and its Commissioners previously commissioned an independent review of the Trust's financial and performance gap in 2017/18 and 2018/19, which produced a set of recommendations for improvement. A key recommendation from this work was to carry out a Demand and Capacity Review to understand the relationship between resources, performance and finances following the introduction of new ambulance performance targets in July 2017, as part of the Ambulance Response Programme (ARP). In line with a number of other ambulance trusts, SECAmb will not be sufficiently resourced to meet these targets by September 2018.

Against this backdrop, SECAmb and its lead Commissioners have jointly commissioned Deloitte and Operational Research in Health (ORH) to undertake this Demand and Capacity Review. The main aim of this review was to agree a timescale for SECAmb to meet national performance standards deemed achievable by SECAmb and Commissioners, estimate the additional resources required (staff, fleet, and finances) to achieve this and to develop a contracting framework to align the incentives of the Trust and its Commissioners, and allow flexibility in the contract if expected assumptions deviate from plan. For the purposes of the review, the operational modelling was undertaken by ORH and financial modelling was undertaken by Deloitte.

As part of the governance of the review, joint working parties were established in the form of governance groups (Strategic Oversight Group (SOG), Steering Group, Focus Sub-Groups) to sign off the key assumptions and analysis, and they were regularly updated on the progress and risks to the review. The outputs of this review have been signed off by SECAmb, Commissioners and the SOG, which includes representatives from NHS England (NHSE) and NHS Improvement (NHSI). Various working papers and documents have been circulated throughout the course of this review, which set out the background to the conclusions summarised in this document. This Final Summary Report is from the Final Report, which sets out further detail on the Demand and Capacity Review.

Scale of change

In 2016/17, SECAmb were not meeting ARP standards, with a Cat 1 mean performance of 08:11 minutes against the national standard of 07:00*, and Cat 4 90th percentile performance significantly worse at c. 300 minutes as compared to the national standard of 180 minutes (converted from the old Red/Green standards based on operational modelling). Consequently, in order to deliver patient and system benefits through improved response times:

- Meeting ARP standards will require a significant improvement to current performance (frontline and Emergency Operations Centre (EOC)**), given this will be the first time in several years that SECAmb will meet standards and this change is estimated to be achieved over an eight month period; and
- A significant uplift in resources (staff and fleet) will be required. The SOG acknowledges this is not achievable in the short term, given the need to:
    - Agree new frontline rosters with unions and staff;
    - Recruit significant numbers of additional staff (frontline and Emergency Operations Centre (EOC)**); and
    - Reconfigure and increase the vehicle fleet to include more ambulances and intermediate tier vehicles (which are more expensive) relative to solo response vehicles (including those led by specialist paramedics), compared to 2016/17 levels

(*) In June and July 2018, SECAmb's Cat 1 mean response time was 07:41 and 08:19 respectively, as compared to a national average of 07:37 in both months (Ambulance Quality Indicators: Systems Indicators). These are in-month estimates and as such are not directly comparably to the 2016/17 average of 08:11 which includes the impact of seasonality.
(**) The impact on EOC is discussed separately

Executive summary (2/8)

Due to capacity constraints in the short term, the SOG considered meeting standards from Q1 2019/20 to be a realistic timeframe

Modelling approach to developing the implementation scenario

Two operational models that varied in terms of frontline skill mix and vehicle composition were considered for comparative purposes, to understand the additional resources required to meet ARP standards. These were the Paramedic Led Ambulance Model, which focused on getting a fast ambulance response on scene, and the Targeted Dispatch Model (TDM), which focused on getting the most clinically appropriate response on scene for the level of acuity.

The SOG recommended that the TDM should be taken forward given i) both models met standards in 2020/21; ii) the TDM provided relatively greater flexibility on using specialist resources as it had a more targeted set of dispatch protocols; iii) the resource requirement for this model was lower, and as such, was likely to have smaller recruitment challenges; iv) it was considered to be more integrated with the wider system; and v) it was estimated to meet standards at a lower cost given the lower resourcing requirement.

While meeting standards from September 2018 was considered in line with national requirements, the SOG agreed that a realistic timeframe for SECAmb to achieve ARP standards would be from Q1 2019/20 taking into account current performance, future demand, rostering, recruitment and vehicle requirements. Therefore, an implementation scenario was developed based on this timeline.

In addition to the above, bottom-up operational modelling was also undertaken for the EOC, to estimate the number of Emergency Call Takers, Clinical Supervisors and Dispatch desk staff required to meet national targets, increase Hear and Treat rates, etc.

Modelling assumptions

The following assumptions underpin the operational and financial modelling undertaken:

Table: Implementation scenario assumptions overview

Description	Q3 2018/19	Q4 2018/19	2019/20	2020/21
Growth	Agreed demand growth of 3.9% per annum			
Inflation	Based on NHS Improvement inflation rates			
Efficiency	Based on achieving a 2% efficiency per year, in addition to operational efficiencies such as re-rostering			
Frontline staff plans	Developed by SECAmb with input from Commissioners and Health Education England (HEE), based on the Full Time Equivalent (FTE) required to meet standards as per operational modelling outputs			
Fleet plans	Based on operational modelling outputs, updated by SECAmb with input from Commissioners, to develop a trajectory that could be operationalised. This includes Double Crewed/Staffed Ambulances (DCAs/DSAs), Intermediate Tier Vehicles (ITVs) and Solo Response Vehicles (SRVs, including those staffed by Critical Care Paramedics (CCPs) and Paramedic Practitioners (PPs))			
Re-rostering	From Q1 2019/20 onwards			

Source: Demand and Capacity review

Executive summary (3/8)

The cost of meeting ARP standards was estimated at between c. £203m to £208m for 2019/20 and c. £215m to £219m for 2020/21

Cost of the implementation scenario

Based on the assumptions set out on the previous page, financial modelling estimated that it would cost SECAmb c. £203m to £208m* in 2019/20 and c. £215m to £219m* in 2020/21 annually to meet national targets in each quarter of the year after accounting for growth. While it is assumed that performance standards will not be met in 2018/19, response times are expected to improve under the modelled scenario throughout the year, which could lead to improved outcomes for patients. Note that for Q3 and Q4 of the current financial year (2018/19) the estimated cost is c. £49.8m to £50.4m* and c. £50.8 to £51.5m* respectively.

Table: Key statistics under the implementation scenario

Category	Metric	2016/17 (actuals)	Q4 2019/20 (modelled)	Q4 2020/21 (modelled)
Performance* (Trust level)	**Cat 1 mean performance (Target: 07:00)**	08:11	06:58	06:58
	Cat 1 90th %ile (Target 15:00)	16:36	14:24	14:42
Staff	**Frontline FTE** (average in-post)	1,734	2,188	2,258
	Weekly paid staff hours (FTE + overtime)**	c. 49k	c. 58k	c. 61k
	Weekly paid staff hours (Private Ambulance Service (PAS))**	c. 7k	c. 7k	c. 5k
Fleet	**DCAs + ITVs**	300	351	363
	SRVs (incl. CCP and PPs)	245	122	143
Estimated annual cost		***£175.5m***	***c. £203m to £208m***	***c. £215m to £219m***

Source: Demand and Capacity review

Modelling outputs

Operational modelling estimated that under the implementation scenario, SECAmb would need c. 2,413 FTE on average in 2020/21 to meet all standards across the year. The recruitment trajectory developed by SECAmb, with input from Commissioners and HEE indicated that c. 462 staff in addition to FTE in Q1 2018/19 could be recruited by Q4 2020/21. As such, even with additional recruitment, it was estimated that additional hours (i.e. above the rostered shifts delivered within FTE contracted working hours) would be needed in the form of overtime and PAS. For fleet, operational modelling estimated that SECAmb would require 50 additional ambulance by Q4 2020/21 as compared to Q1 2018/19.

A number of tests of reasonableness were undertaken around the implementation scenario estimates, including:

- Testing whether the implementation scenario was sufficiently stretching. As this will require SECAmb to go through significant operational change to meet national performance standards, the implementation scenario was considered stretching, but feasible as the resourcing trajectories were developed on a bottom-up basis.
- Testing whether the scale of change is reasonable in comparison to other Trusts. Financial modelling estimates that SECAmb requires a c. 5%-6% increase in cost per year up to 2020/21. This is similar to the quantum of change at East of England Ambulance Service[1] (EEAST), which required an increase in cost base of c. 5% per annum, however, in the context of a different operational and demographic mix.

(1) https://www.eastamb.nhs.uk/EEAST-ISR-Report-March-2018.pdf
(*) Commissioners and SECAmb need to agree on whether some recurrent costs related to education and training and back-office are to be included within the baseline and costs going forward. The lower bound of the cost estimates is if these recurrent costs are excluded, and the upper bound is if these are included
(**) On average across the year

Executive summary (4/8)

The modelled implementation trajectory estimated that performance against all standards is likely to improve going forward

Frontline performance

The table below shows the Trust level performance trajectory as modelled in the implementation scenario. It can be seen that all ARP standards are met on a quarterly basis from Q1 2019/20, providing greater benefits to both patients and the system. Performance also improves for all standards for most CCGs, although there is variation in performance due to underlying factors such as demography and rurality.

Table: Modelled frontline performance trajectory

Period	Cat 1		Cat 2		Cat 3	Cat 4
	Mean	90th %ile	Mean	90th %ile	90th %ile	90th %ile
Q3 2018/19*	08:31	17:12	15:49	30:42	193:36	308:30
Q4 2018/19*	07:35	15:18	14:53	29:54	58:42	102:18
Q1 2019/20	07:00	14:30	12:41	26:12	51:18	84:36
Q2 2019/20	07:00	14:24	12:49	26:36	55:36	111:36
Q3 2019/20	07:00	14:30	13:08	27:12	67:30	163:36
Q4 2019/20	06:58	14:24	13:33	27:42	65:18	158:06
Q1 2020/21	07:00	14:42	10:23	18:36	73:48	93:48
Q2 2020/21	06:59	14:42	10:32	18:48	89:48	108:00
Q3 2020/21	06:58	14:36	10:33	18:48	118:00	149:24
Q4 2020/21	06:58	14:42	10:41	19:06	117:48	153:00
Target	**07:00**	**15:00**	**18:00**	**40:00**	**120:00**	**180:00**

Source: Demand and Capacity review

This table shows that when Cat 1 standards are met, performance on Cat 2 to Cat 4 standards also improves significantly. This is because Cat 1 standards have a more stringent resourcing requirement as compared to the other standards, so when resources are deployed to meet Cat 1 standards, other standards may over-perform as a result. Further, performance is also contingent on the mix of staff and vehicles deployed. Changes to the staffing and fleet mix as per the workforce plan and fleet plan lead to peaks and troughs in performance estimates over time, in addition to the impact of seasonality (e.g. winter pressures).

EOC performance

In addition to changes to frontline staff, changes to the EOC have also been incorporated to meet the growth in demand and projected performance improvements. The operational model assumes the following with respect to EOC performance:

- **EOC staff:** The model assumed an increase from 308 FTE (2016/17 average) to 398 FTE by Q4 2020/21 for Emergency Medical Advisor (EMA) Call Handlers, Clinical Supervisors and Dispatch FTE that were modelled bottom-up.
- **EMA Call Takers**: The model assumes that 95% of calls are answered within five seconds (in line with the national standard compared to 80.2% currently) from Q1 2019/20 onwards.
- **Clinical supervisors**: It was assumed that a 10% Hear and Treat rate is achieved from Q1 2019/20, with a 40% success rate for a clinician taking a Hear and Treat call
- **Dispatch desk**. New dispatch desk boundaries were also identified to deploy resources optimally.

(*)
1. < Redacted >
2. In June and July 2018, SECAmb's Cat 1 mean response time was 07:41 and 08:19 respectively, as compared to a national average of 07:37 in both months (Ambulance Quality Indicators: Systems Indicators). These are in-month estimates and as such are not directly comparably to the 2016/17 average of 08:11 which includes the impact of seasonality

Executive summary (5/8)

By improving SECAmb's response time performance, the service for patients is expected to improve

Opportunities

The implementation scenario is expected to provide a number of opportunities, including:

- **Improving care for patients:** Response times are estimated to improve from the outset as ARP is implemented and as staff are on-boarded in line with the estimated ARP recruitment trajectory.
- **Benefiting frontline staff**: The staff survey undertaken within the University of Sheffield evaluation report[1] showed that staff viewed the ARP programme as a positive development.
- **For SECAmb** to become operationally and financially sustainable for the long term, whilst meeting national performance standards and supporting the wider system, as the model is built based on the estimated cost of efficient rosters designed to deliver the new ARP standards.
- **For Commissioners** to have greater certainty around their expenditure and service performance, as the modelling is associated with a cost and a simulation-based estimation of performance. Under the implementation scenario, it is estimated that performance improves for all standards for most CCGs by Q4 2020/21.

Issues and risks

Whilst there are a number of opportunities within the implementation scenario, there are also a set of risks associated with it. These include the following. Please note that the RAG rating below is based on a qualitative assessment of the materiality of the risk being realised on the implementation trajectory.

Risk	RAG rating
Additional cost. The service cost in 2019/20 and 2020/21 is greater than in 2017/18 and 2018/19.	
National timelines. Achieving national standards in Q1 2019/20 is later than the national requirement of September 2018.	
Performance variation. While standards are expected to be met across the region as a whole, there will be variation at a local level, driven by factors such as demographics, rurality and health need.	N/A
Recruitment. There is a risk that SECAmb is unable to recruit and train staff in line with the workforce plan agreed by HEE to meet standards.	
Re-rostering. New rosters need to be agreed with Unions and staff, and be implemented by Q1 2019/20. Any delay or change to this could impact upon performance and service cost.	
PAS availability. There may be risks around the availability of PAS due to quality and governance issues around their use. Any contraction in the supply of PAS will impact the available resources and therefore the modelled trajectory.	
Other areas. The implementation scenario is based on a set of assumptions (demand, handovers etc.) that were agreed at a point in time. If these were to be revised in the future then this would impact the estimates set out.	

(1) https://www.sheffield.ac.uk/polopoly_fs/1.716080!/file/ARPExecSummary.pdf

Key: *High* *Medium* *Low*

Executive summary (6/8)

The contracting mechanism sets out how SECAmb could be measured and reimbursed for 999 services based on the output of this review

Contracting mechanism

The contract formalises how the resourcing and financial implications are paid for. This includes the mechanism through which the Trust is paid, as well as how local areas split the contract value between them in order to reflect their geographic differences.

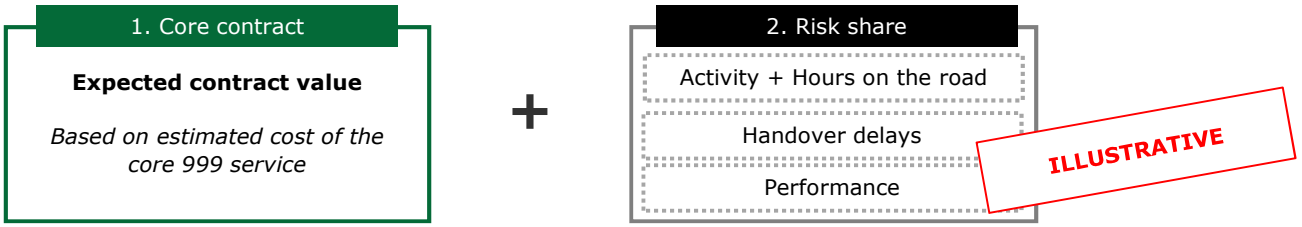
The SOG agreed that the 2018/19 contract variation will be agreed between SECAmb and Commissioners outside the Demand and Capacity review. Cost estimates for Q3 (c. £49.8m-£50.4m) and Q4 2018/19 (c. £50.8m-£51.5m) as per the review could be used to inform the contract value for the remainder of the year. Discussions to date suggest that a block contract is being considered between SECAmb and Commissioners for the remainder of 2018/19. Please note that this report does not reflect the current stage of contract negotiations. Commissioners and SECAmb noted that the adjustment mechanisms set out below could also be considered for the 2018/19 contract variation. It should be noted that since 2018/19 is a transitional year where standards will not be met, the choice of levers should be carefully considered.

The expected value of the contract in 2019/20 and 2020/21 could be based on SECAmb's estimated future demand and required capacity to deliver the modelled performance trajectory. As noted earlier, operational and financial modelling assumed a demand growth of 3.9% per annum up to 2020/21 and incorporated efficiencies around rostering, increased Hear and Treat rates, along with an efficiency envelope based on 2% savings per annum, in line with the national efficiency requirements. This was deemed a sensible assumption by SECAmb and Commissioners and does not reflect a specific analysis of schemes underpinning it.

Discussions with SECAmb and Commissioners suggested that the contract could comprise an expected 'core' contract value, with a risk share on the following four potential drivers of cost:

- **Activity** (by currency). SECAmb currently has marginal rates in place to adjust the contract value if activity deviates from plan (either above or below). Activity deviation from plan could be reimbursed through marginal rates.
- **Hours on the road**. Hours put out on the road could act as a binding constraint on whether the activity adjustment is used, so that if hours go down when activity goes up, the contract value could either not change, or could imply a transfer of funds back to the Commissioners from SECAmb.
- **Handover delays**. Deviations in handover hours lost could trigger year-end changes to the contract value based on the relationship between handover delays with cost.
- **Performance**. After accounting for changes in performance due to handover delays, remaining decreases in performance could adjust the contract value.

Contracting Mechanism 2019/20 – 2020/21



Executive summary (7/8)

The current contract split is based on CCG level activity, however, other potential options are available

Funding the contract

Currently, Commissioners fund the contract based on their share in the total activity. Three potential options have been set below; these are not exhaustive and there are others which could be considered.

Commissioners will agree the preferred approach to split the contract between themselves outside of the Demand and Capacity review.

Funding the contract

Three example approaches for splitting the contract across Commissioners are set out below:

1. **Activity**

This approach estimates the contract shares based on the CCG's share in the total activity at a currency level.

2. **Activity + Resource**

This approach estimates the contract shares based on activity and the resources used to service that activity within CCGs at a currency level. This would capture local CCG growth assumptions and could be adjusted for acuity by differentiating this by call type (e.g. Category 1,2,3,4). For 'See' activity, average job cycle time (JCT) is used to estimate the total minutes spent in each CCG area.

3. **Activity + Resource + Performance**

This approach adjusts approach two for performance, by adjusting contract shares based on the difference from the Cat 1 mean national response time standard at a CCG level. A more complex approach, such as developing a composite index across all national performance standards, could instead be developed.

ILLUSTRATIVE

Executive summary (8/8)

The key next steps are set out below

Conclusions

This review has highlighted that significant investment is required to enable SECAmb to meet its future demand growth and comply with ARP standards within a reasonable timescale. Notwithstanding the scale of investment, significant benefits are expected to flow as a consequence of improving SECAmb's performance, including:

- **For Patients:** Improved quality and reliability of service matched against their acuity;
- **For SECAmb:** Longer term future proofing of its business, improved staff morale and transparency over how its performance will be monitored and paid for;
- **For Commissioners:** An improved, compliant ambulance service, and flexibility in the contract management linked to performance, providing greater understanding, clarity and transparency over cost of the service being provided.

Next steps

The key next steps are set out below:

- Further socialisation of the outputs of the Demand and Capacity Review by lead Commissioners with other stakeholders (e.g. Associate Commissioners, STP leads etc.)
- SECAmb and Commissioners to develop a Service Development and Improvement Plan (SDIP), if required
- SECAmb and Commissioners to continue to work together to further develop the clinical case for change and the ability to articulate clinical management of risk, including clinical and patient outcomes and impact
- SECAmb to further develop an implementation plan for meeting the ARP standards (e.g. the recruitment and fleet trajectories), and carry out on-going engagement with Commissioners and HEE on these plans
- Wider monitoring in line with the national Ambulance Improvement Programme to continue
- Continued monitoring of performance in line the national Ambulance Quality Indicators (AQIs), which comprise the Systems Indicators (AmbSYS) and the Clinical Outcomes (AmbCO)
- SECAmb / Commissioners to identify and implement programmes to achieve efficiencies
- SECAmb / Commissioners to agree and draft the terms of the contract, for Q3 and Q4 of 2018/19 and 2019/20 and 2020/21
- Commissioners to agree a method for splitting the contract
- SECAmb / Commissioners to monitor and track differences between expected and actual performance on key contract metrics
- SECAmb / Commissioners could develop a dynamic modelling tool to inform future year contracting based on the outputs of the Demand and Capacity review, and the estimated relationships between resource, performance, and finances

Deloitte.