UNIVERSITY of STIRLING

The Scottish Ambulance Service

New Clinical Response Model

Evaluation Report

Authors: Dr Kathleen Stoddart (Principal Investigator)
Dr Julie Cowie
Dr Tony Robertson
Associate Professor Dr Carol Bugge
Professor Jayne Donaldson
Dr Federico Andreis

Date Published: February 2019

Table of Contents

An *addendum* provided by the Scottish Ambulance Service is included *verbatim* (page 62)

Acknowledgements

The Scottish Ambulance Service Information Services for collating and providing summary data. Richard Dobbie and his team at NHS Information Services Division who collated data and contributed to analysis have also supported the evaluation.

Glossary of key terms

AMPDS	AMPDS Advanced Medical Priority Dispatch System
CAD	Computer-Aided Dispatch
DCR	Dispatch Code Reference
CI	Confidence Interval
ED	Emergency Department
GROS	General Register Office for Scotland
ILT	Immediately Life Threatening
ICD	International Classification of Disease
ISD	NHS Information Services Division
MPDS	Medical Priority Dispatch System
NCRM	New Clinical Response Model
NRS	National Records for Scotland
PDA	Pre-Determined Attendance
SAS	Scottish Ambulance Service
SD	Standard Deviation
UCDM	Unscheduled Care Data Mart

Executive Summary

The Scottish Ambulance Service (SAS) responds to around 1.8 million calls per year, including responses to 700,000 emergency and unscheduled incidents. Of these responses, over 500,000 are received through the 999 call service. SAS transfers around 90,000 patients between hospitals each year and responds to over 150,000 urgent requests for admission, transfer and discharge from GPs and hospitals (SAS, 2015). In 2017 SAS began to implement a new clinical response model (NCRM). The **aims of the NCRM** are to:

- Save more lives by more accurately identifying patients with immediately life-threatening conditions, such as cardiac arrest;
- Safely and more effectively send a matched resource first time to all patients based on their clinical need.

The University of Stirling, commissioned to carry out an independent evaluation of the NCRM using data provided by SAS and NHS Information Services Division (ISD), considered the following questions:

1. *Are patients with Immediately Life Threatening (ILT) conditions more quickly and accurately identified?*
2. *Are more lives saved as a consequence of the best available resources being dispatched to the patient?*
3. *Are improved clinical outcomes achieved if the matched resources are sent first time for patients with non-ILT conditions?*

Methods

A quantitative analysis was conducted comparing SAS data on response to 999 calls from a pre-NCRM implementation time-period (January 2016) and a post-implementation time-period (January 2017 and January 2018). NHS ISD linked additional data from the Unscheduled Care Data Mart (UCDM) to the SAS data. UCDM contains emergency department data (ED) and data from the National Records of Scotland (NRS) for mortality data.

Data were examined for the purple code (the highest risk category of call to the 999 service) and within the purple category, those patients in cardiac arrest. The same analyses were conducted for the remaining colour codes and a selection of clinical categories within these colour codes: breathing difficulties (red), stroke (amber) and falls (yellow).

Key Findings

Interpreting this data

It should be noted that data is taken from only three (and in some cases two) time points and only from the month of January. While this does allow some relevant comparisons between the years, the findings cannot be generalised to the whole year and the whole time-period in question (January 2016 – January 2018). In addition, call volume was approximately 9% higher in 2018 compared to 2017 and 2016 (which were similar) with over 4,000 more calls in January 2018. Further analysis of the data using data from each month, as well as individual-level data (rather than it being aggregated), would allow much more robust and relevant evidence of change and the impact on the service and patients.

1. *Are patients with Immediately Life Threatening (ILT) conditions more quickly and accurately identified?*

Patients with ILT conditions (purple calls) would appear to be more accurately identified post-NCRM with a noticeable increase in patients coded with ILT conditions by 2018. The time to respond to ILT conditions was slightly longer (but not statistically significant).

Speed

Resource allocation was used as an indicator of speed of identification. We found that resource allocation (and in turn response times) did not differ significantly between January 2016 (pre-NCRM) and January 2017 (post-NCRM introduction) for ILT (purple) calls. However, there was a longer time to allocate resources (i.e. identify) purple calls in 2018 compared to 2016 and this was statistically significant. For all other colour codes, 2017 and 2018 resource allocation were also significantly slower than 2016 (except amber 2017 calls) as expected with a priority-based system.

Call handlers were provided with further training and development in the process of triage over the course of 2016 onwards, with the aim of more accurately allocating patients into the most appropriate category, and therefore it was to be expected that time to allocate resources and identification into the correct category would take longer.

Accuracy

Comparing 2016 (pre-NCRM) and 2017 (post-NCRM introduction) outcomes data, we found that sensitivity (correctly identifying a purple, ILT condition) was higher in 2017 compared to 2016, but specificity (correctly identifying a non-ILT condition) was lower in 2017. Overall accuracy (the likelihood of being correctly identified as either ILT or non-ILT) was not different between the two-time points. Similar results were also seen for the cardiac arrest cases within the purple calls.

2. *Are more lives saved as a consequence of the best available resources being dispatched to the patient?*

Survival for purple-coded patients is markedly lower with respect to all other causes (as one would expect) and reflects that purple-coded calls/conditions are a unique category (in terms of risk of death) and represent the majority of incidents where patients face an immediate threat to life (ILT). The risk of death across the other colour codes is small in comparison and therefore differences of survival seem to exist only for the purple-coded patients.

The cardiac arrest rate within the purple coded is around 53%. Survival analysis for all patients within the purple code and specifically for those affected by cardiac arrest are considered next.

There seems to be a considerable (~20%) increase in survival for all purple-coded patients comparing January 2016 to January 2017, which is constant over time from time 0 (confirmed dead when the ambulance arrives at the scene) to 30 days post-call. When comparing January 2016 to January 2018 for the same group, survival also increased (~10%).

The number of lives saved, 30 days post-call, in patients with ILT conditions in January 2016 (pre-NCRM) was 32 (14.2% of purple calls), and in post-NCRM in January 2017 was 134 (28.6% of purple calls) and in January 2018 was 182 (26.6% of purple calls).

Although the numbers of patients with ILT conditions has increased, the data from the specificity and sensitivity analysis (Table 14) shows that there is no difference in false positive rates between the years. This suggests that the acuity of these patients remains very high and that the increase in volume represents patients correctly identified with the highest requirement for immediate response. Therefore, the increase in survival probability with those with ILT conditions is not likely to be caused by artificial inflation caused by conservative allocation of patients with ILT conditions to the purple code but rather by appropriate allocation and intervention(s) to those patients at risk from death due to ILT conditions. In terms of the 2018 survival probability being lower than in 2017, it is possible that the higher call-load in 2018 has limited the impact previously seen in 2017. Continued monitoring of these data is needed to identify how mortality has been impacted by the NCRM over the longer-term.

3. *Are improved clinical outcomes achieved if the matched resources are sent first time for patients with non-ILT conditions?*

Overall survival for all non-ILT codes (Red, Amber, Yellow and Green) was similar, as noted above (where purple calls carry much higher risk of death). For these codes there was also no clear difference in survival in 2017 versus 2016 or 2018 versus 2016.

Breathing difficulty (a sub-set of the red calls) seems to have worsened between 2016 and 2017, with 451 patients having a decrease in survival from 3% to 6%, with the gap widening as time passes. However, by 2018, survival was at 2016 levels despite the number of incidents (n=2044) back to the levels seen in 2016 (n=2018). No differences between years seem to be present for stroke or falls.

Data on further clinical outcomes were not available within this dataset to analyse in any further detail.

Conclusions

By January 2018 the number of incidents (n=52,871) had increased by 9% when compared to January 2016 (n=48,544), amounting to over 4000 more incidents in 2018 than seen in 2016 or 2017. During this time of high demand in 2017 and particularly 2018, the NCRM does accurately identify patients who have the greatest need for services from SAS. The NCRM's identification and triage of patients into triage categories, although taking time for the call handler and dispatching system, can get the ambulance and its crew to patients with the greatest need and this has improved the survival of those with immediate life-threatening conditions. Those with lower acuity needs are responded to but in a longer time period as expected when using a priority-based system (but with no apparent impact on survival). These conclusions are reached in the context of analysing aggregated data over three fairly short time-periods and further research over a longer time frame, with longitudinal data on individual cases, would further improve the evidence base for the NCRM.

Chapter 1: Background and context in Scotland

The Scottish Ambulance Service (SAS) conducted an extensive internal review of over 500,000 individual 999 patient incidents over one calendar year, demonstrating that some patient conditions were not accurately identified as being Immediately Life Threatening (ILT) and non-ILT. This inaccuracy was determined to be a risk to patients with impacts on both the response and the resources dispatched by SAS. Consequently, a new clinical response model (NCRM) was developed. The model it replaced had been in place since the mid-1970s with few changes over time. The old performance framework, with outdated targets, is dissonant with contemporary pre-hospital and in-hospital models of health care. That concern is echoed in a growing body of literature about the development of new response models and their impacts (Blackwell et al 2009; Bigham et al 2013; Turner et al 2015; O'Neill et al 2017). The overarching concern was that applying the response and performance framework devised in the 1970s no longer delivered the most effective and efficient clinical care to patients.

The old performance model focussed on response time of eight minutes for everyone and led to the development of well-intentioned, but ultimately unhelpful practices. The SAS identified these unhelpful practices as including:

- dispatching resources with no awareness of the patient's clinical need (dispatch on address);
- dispatching multiple resources in the hope that one would hit the target
- mismatching the type of resource to the need of the patient, for example sending an ambulance car which would hit the target but would result in a delay in the patients' ultimate conveyance to definitive care.

In November 2016, SAS began to pilot the NCRM for emergency 999 calls with the ambition to help save more lives and to match response to patient need more effectively. SAS aimed to take more care to the patient and, where it is in the best interests of the patient, enable more care provision within communities rather than conveyance to an Emergency Department (ED) and admission to hospital.

The NCRM included a change from the previous A, B and C categories to a four-response category grouping: purple, red, amber and yellow. The purple, red, amber and yellow categories relate to the level of clinical acuity with triage-determined response levels. The new categories are based on an internal SAS review of ~500,000 incidents and analysis of interventions and outcomes experienced by patients across the whole 999 triage system. For example, patients assigned to the purple category have higher levels of immediate interventions in terms of breathing support and have a cardiac arrest rate over 10%. Both purple and red calls still have the historical eight-minute target applied to them, however a move to reporting median and 90th centile response times, which SAS views as a more meaningful description of response performance, is also now in place.

It was the view of SAS's senior clinicians that the old performance model did not reflect the changing expectations of patients, SAS clinicians or the wider health and care system. A new model needed much more than the singular expectation of conveyance to the nearest ED, as was the driver in the previous time-based model. Considering the increasing skill set of SAS clinicians aligned to improved diagnostic equipment and treatment options, as well as other developments[1] in Scotland, there was a requirement to produce a more clinically-led response to how 999 calls are triaged, and to how the most appropriate resource is despatched to meet patients' needs.

Response categories

Based on clinical evidence form a review of ~500,000 cases, including SAS interventions and conveyance decisions, the response categories are (Figure 1):

Purple – for a code to be in this category there needs to be a cardiac arrest rate for patients of >10%. In fact, the cardiac arrest rate in this category is >50%.

Red – For a code to be in this category, the cardiac arrest rate for patients is between 1% and 10%. In fact, the mean rate is 1.3%.

[1] Scotland's Out of Hospital Cardiac Arrest strategy, the Scottish Patient Safety Program work about identification of deteriorating patients (for example. sepsis), definitive care pathways for patients affected by conditions such as stroke and heart attack, an expectation to shift the balance of care and take more care to the patient and Scotland's, at that time emerging Trauma Network plans in 2016

SAS describe both purple and red categories as ILT, however given the significant differences in cardiac arrest rates and volumes between these codes, the Stirling report considers purple codes as ILT for the purposes of their analysis.

Amber – patients in this category have a cardiac arrest rate of less than 1%, but a high likelihood of need for definitive care. For example, patients affected by stroke who need to go to a specialist stroke unit or patient with chest pain who may be experiencing a heart attack and need to go to a primary percutaneous coronary intervention centre for emergency intervention.

Yellow – these patients require an emergency ambulance response but have a range of conditions that are not life threatening and whose care requirements can be met either in a local Emergency Department or by treatment at scene.

Green – all remaining patients are included in this category.

SAS senior clinicians expected that via the NCRM there would be changes in categorisation that would result in an increased number of patients in the purple and amber categories, a stable number of patients in the yellow category and a significant reduction in the red (previous Cat A) category as outlined in Figure 2. Under the old model, purple codes were responded to under the red (Category A) group and amber codes applied only to certain stroke patients and fell within the yellow (Category B) group. All other cases were deemed as green/Category C.

New Clinical Response Model Hierarchy With Outline Principles
Purpose Response Category ▪ Highest response priority ▪ Cardiac arrest rate over 10% ▪ Respond with closest resource ▪ Paramedic attendance essential ▪ Minimal of three responders to scene + double crewed ambulance if not in that response ▪ Consider partner agencies to support response
Red Response Category ▪ Second highest response priority ▪ Cardiac arrest rate >!% and defined need for resuscitation ▪ Response with the closest resource + double crewed ambulance if not that response ▪ Paramedic attendance essential
Amber Response Category ▪ Third response category ▪ <1% cardiac arrest rate ▪ Defined need for acute pathway care ▪ Response with the right resource – emergency transporting ambulance ▪ Paramedic attendance preferred
Yellow Response Category ▪ Fourth response category ▪ <1% cardiac arrest and no defined acute pathway care ▪ Response with the right resource – ambulance for defined hospital need and PRU for potential alternative pathway care
Green Response Category ▪ Fifth response category ▪ Exclusion of above categories ▪ Potential for additional clinician led telephone triage ▪ Face to face assessment when required

Figure 1 New Clinical Response Model Colour Categories

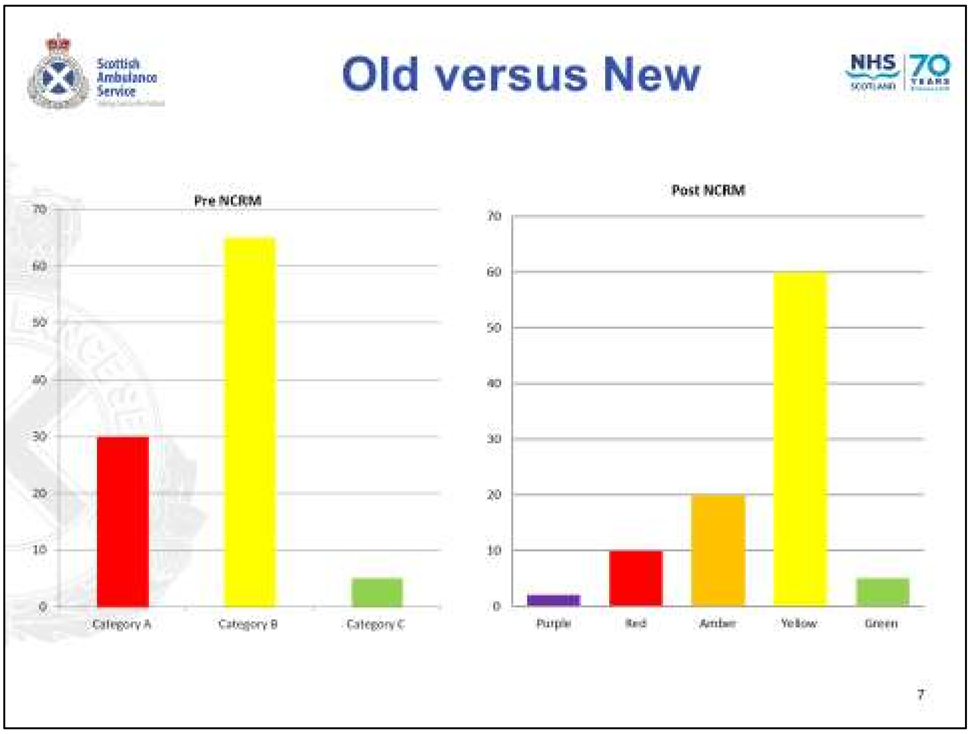


Figure 2 Predicted Number of Colour Categories between Old and New Clinical Response Model

Medical Priority Dispatch System

The Medical Priority Dispatch System (MPDS) is populated by the colour categories described above, with associated triage-determined response levels. For example, purple denotes a high risk of cardiac arrest, within a larger ILT category, two ambulances crewed by paramedics would be dispatched immediately.

The response time targets prior to the NCRM were as follows:

- 80% of Cardiac Arrest calls - 8 minutes
- 75% of CAT A (Red and Purple) - 8 minutes
- 95% of CAT B (Amber and Yellow) - 19 minutes
- 91% of 1 Hour urgent GP referrals - within 1 hour

The time targets have remained the same since the introduction of the NCRM:

- Purple - 8 minutes
- Red - 8 minutes
- Amber - 19 minutes
- Yellow - 19 minutes

Dispatch on disposition was introduced in October 2017, where the assessment and management of 999 calls allows for the allocation of an appropriate resource when disposition is reached. In other words, when the triage system has reached the 'final dispatch code' and thereby determined the priority of that call. Previously dispatch on address was the procedure followed, whereby if an ambulance resource was available then it would be dispatched as soon as the address was verified, but before any clinical information was available. This change to dispatch on disposition was also combined with SAS now reporting against median and 90th percentiles for ILT calls as follows:

- Median Purple Target - 6 minutes
- Median Red Target - 7 minutes
- 90th Percentile ILT (Red and Purple) - 15 minutes

Aims of NCRM

The aims of the NCRM are to:

- Save more lives by more accurately identifying patients with immediately life-threatening conditions, such as cardiac arrest
- Safely and more effectively send a matched resource first time to all patients based on their clinical need

Whole System Approach

The whole system approach to implementation meant that the NCRM, introduced in two phases, commenced in November 2016, involving the application of a revised triage model, based on clinical data, as described above. Other system and training initiatives were included, for example:

- Training around the new response model for all call handlers
- Introduction of ELAN telephone interface, allowing automatic address verification from land lines
- Upgrade to mapping of Public Accessible Defibrillators onto the SAS IT system.

In Phase 1 there was no change to the practice of dispatching on address. Moving away from dispatch on address was part of Phase 2. In order to move away from dispatch on address, the triage system identified patients in cardiac arrest rapidly. This involved the development of Pre-Entry Questions (PEQs) which identified patients who were not breathing or unconscious before formal triage started and dispatched to these patients as a 'purple' response. The data analysis to support the introduction of 'Pre-entry Questions' (PEQ) (to identify patients in, or at high risk of, cardiac arrest before they enter the formal triage system) was completed in the summer of 2017. Prior to changing the dispatch model, the first PEQ was introduced to test this in the live environment before making any changes to the historical model of dispatching on address verification. Phase 2, introduced in October 2017, involved the addition of a second PEQ, and the move to 'Dispatch on Disposition'. This means that unless a response is triggered by PEQ, the resource is not dispatched until a final triage code is reached. The aim here is for rapid identification of patients in cardiac arrest, or at high risk of cardiac arrest, with SAS then able to match the most appropriate available resource to patient need for all other 999 calls.

Phase 1 – exploration of changes to the triage of calls. The 999 call-handling processes were modified to incorporate questions that immediately identify the most urgent calls. For all other calls additional time is allowed to support targeting the right resource to the right patient and reducing allocation of multiple resources in order to 'stop the clock'. From SAS specifications, the steps in phase 1 included:

- Perform clinical evaluation of clinical measures and outcomes
- Define, write and agree corporate communication plan for stakeholders
- Re-write the Dispatch Code Reference (DCR) response table to improve response times for ILT patients
- Ensure Computer-Aided Dispatch (CAD) colour categories align with the new response model
- Input predetermined attendance (PDA) type into dispatch cross reference (DCR) table aligned with MPDS determinant
- Plan and deliver training/awareness to agenda for change and operational staff
- Review and amend real time reporting system and data warehouse reports
- Obtain formal clinical sign-off for MPDS elements by the Clinical Advisory Group;
- Define, write and agree staff engagement plan
- As a result of the increasing number of Public Access Defibrillators (PAD) being installed at various locations throughout Scotland, it was necessary to activate the Enhanced Defib Module to electronically manage their use

- Install the C3 Nexus ELAN Module in order to automatically identify the location of 999 callers from landlines.
- Introduce Paramount v5.1 Breathing Detector Tool which allows call handlers to identify the breathing rate of the patient
- Introduce the use of the Paramount SHIFT feature to align with DCR priorities

Phase 2 - a continuous review of call categories and development of a new set of categories that align clinical and resource allocation requirements and response options for 999 dispatch codes. Key activities in phase 2 as specified by SAS were:

- Test and introduce Pre-Entry Questions, in order to identify patient in, or at high risk of cardiac arrest earlier in the process
- Activate the Enhanced Dispatch Module to effect 'Dispatch on Disposition' in order to change the current process of dispatching to all incidents irrespective of whether the call is high priority, at the earliest opportunity, including the 'Front Loaded Question' (FLQ)
- Activate the First Responder Model (Whiteboard – Static FR) and (Smartphone – Mobile FR), increasing the volume of calls responded to by Community First Responders
- Introduce Paramount MPDS v13, which brings a number of significant benefits including the facilitation of quicker identification of cardiac arrests
- Activate the Key Phrases Module to provide an early indication of the likelihood of a call being classed as immediately life threatening
- Work towards an increase in referral calls (hear and treat) for those patients for whom and ambulance response is not appropriate.

An *addendum* provided by SAS is included at the end of this repot (page 63). The areas addressed in the addendum provide more information of the NCRM and are 'Scottish Ambulance Service – New Clinical Response Model Clinical Development Overview; New Clinical Response Model Phases; Call Volume and Clock Start'.

This Chapter provided background information to the whole system of change involved in the implementation of NCRM. The focus of the evaluation was on the questions posed and did not evaluate parts of the 'system' nor the experiences of staff and/or patients.

Chapter 2: Approach to Independent Evaluation

The University of Stirling, commissioned to carry out an independent evaluation of the NCRM using data provided by SAS and NHS Information Services Division (ISD), considered the following questions:

1. *Are patients with Immediately Life Threatening (ILT) conditions more quickly and accurately identified?*
2. *Are more lives saved as a consequence of the best available resources being dispatched to the patient?*
3. *Are improved clinical outcomes achieved if the matched resources are sent first time for patients with non-ILT conditions?*

These questions align with the aims of the new model.

A quantitative analysis was conducted comparing SAS data on response to 999 calls from a pre-NCRM implementation time-period (January 2016) and a post-implementation time-period (January 2017 and January 2018). NHS ISD linked additional data from the Unscheduled Care Data Mart (UCDM) to the SAS data. UCDM contains emergency department data (ED) and data from the National Records of Scotland (NRS) for mortality data. Data were provided for three time points: January 2016 (01/01/2016-31/01/2016), January 2017 (01/01/2017-31/01/2017) and January 2018 (01/01/2018-31/01/2018).

Data were examined for the purple code (the highest risk category of call to the 99 service) and within the purple category, those patients in cardiac arrest. The same analyses were conducted for the remaining colour codes and a selection of clinical categories within these colour codes: breathing difficulties (red), stroke (amber) and falls (yellow).

A quantitative analysis of trends for a broad range of operational performance and resource utilisation indicators within the old model and the NCRM. The specific analysis methods are detailed below in the relevant sub-sections.

It was not possible for individual patient-level data to be shared for this evaluation, instead summary-level data in the form of counts of calls or mean times for each colour category or condition-specific category were available for 2016, 2017 and 2018 ('SAS data'). 2016 and 2017 data (but not 2018 data) were available for patient outcomes provided and linked to patient records by ISD ('ISD outcomes data'). For the 'outcomes data' ISD selected the last SAS event in the pathway and indexed off from that point in order to measure outcomes according to the remaining sequence of events.

Initial descriptive analysis of the 'SAS data' revealed that there were no differences in the overall number of incidents recorded between 2016 and 2017, with 48,544 in January 2016 and 48,588 in January 2017 (equivalent to a 0.09% increase from 2016 to 2017), confirming that these time points represent a valid comparison (at least in terms of overall incidents). However, in January 2018 there were 52,871 incidents, an almost 9% increase compared to the same month in 2016 and 2017. The reasons for this change have not been explored in this report as this was beyond its scope.

Chapter 3: Analyses

A series of descriptive and inferential analyses were performed on the 'SAS data' and 'outcomes data' in order to answer the key research questions posed in Section 2. These are outlined below (Sections 3.1-3.3).

3.1 Descriptive Analyses

3.1.1 Overall Incidents

Although there were no differences in the overall number of incidents between 2016 and 2017, there were differences when broken down by colour code. Purple, amber, yellow and green codes were all higher in 2017 compared to 2016. In contrast, red incidents were lower in 2017.

This pattern continues in 2018 (Table 1). Incidents categorised as purple were higher in 2017 vs 2016 with a 70% increase, although as a proportion, the total number of incidents remains relatively low in this group in both years (0.9% in 2016 and 1.5% in 2017). These increases continued in 2018, resulting in a 121% increase in purple incidents compared to 2016, although this represents only a 0.2% increase in terms of purple incidents as a proportion of all incidents compared to 2017.

Red incidents were lower in 2017, reducing from 12,847 in January 2016 (26.5% of all incidents) to 4,167 in January 2017 (8.6% of all incidents. Red incidents remained at a similar level in 2018 (10.3% of all incidents).

Amber incidents showed a more than four-fold increase in 2017 and these continued to increase in 2018, now representing 22% of all incidents (compared to 3.2% in 2016). There were smaller (2-5%) increases between 2016 and 2017/18 for Yellow and Green categories (Table 1).

Table 1 Number of incidents in January 2016, 2017 and 2018 and changes between years

Incident Category	January 2016 (% versus other codes)	January 2017 (% versus other codes)	January 2018 (% versus other codes)	2016 – 2017 Absolute Change	2016 – 2017 % Change	2016 – 2018 Absolute Change	2016 – 2018 % Change
Purple	414 (0.9%)	704 (1.5%)	916 (1.7%)	+290	+70.05%	+502	+121.26%
Red	12,847 (26.5%)	4,167 (8.6%)	5,460 (10.3%)	-8,680	-67.56%	-7,387	-57.50%
Amber	1,554 (3.2%)	8,930 (18.4%)	11,623 (22.0%)	+7,376	+474.65%	+10,069	+647.94%
Yellow	28,184 (58.1%)	28,879 (59.4%)	28,531 (54.0%)	+695	+2.47%	+347	+1.23%
Green	3,791 (7.8%)	3,997 (8.2%)	3,973 (7.5%)	+206	+5.43%	+182	+4.80%
Unknown	1,754 (3.6%)	1,911 (3.9%)	2,368 (3.6%)	+157	+8.95%	+614	+35.01%
Total	**48,544 (100.0%)**	**48,588 (100.0%)**	**52,871 (100.0%)**	**+44**	**+0.09%**	**+4,327**	**+8.91%**

3.1.2 Purple Incidents

The number of purple incidents allocated and mobilised in January 2017 was 703, a 70.2% increase compared to the same period in 2016. This change mirrors the overall incident comparisons described above and in Table 2 ('Incidents with Resources'). All resources in the purple category that were allocated were mobilised in both 2016 and 2017. Although the numbers of resources that arrived at the scene was lower compared to those mobilised (and allocated), this drop was relatively small in both 2016 (413 vs. 407) and 2017 (703 vs. 691). Therefore, although more incidents were classed as purple in 2017, this did not appear to change how purple-coded incidents were handled in terms of resources being mobilised and arriving on scene. For each incident, however, there may be more than a single resource allocated, mobilised and that arrives on scene (for example, an ambulance and a paramedic response vehicle ('Total Number of Resources', Table 2).

Table 2 January 2016 and 2017 resource allocation, mobilisation and arrival at scene for purple-coded incidents

Resources	January 2016 (% versus other codes)	January 2017 (% versus other codes)	2016 – 2017 Absolute Change	2016 – 2017 % Change
Incidents	414 (0.9%)	704 (1.4%)	+290	+70.05%
Incidents with Resources Allocated	413 (0.9%)	703 (1.5%)	+290	+70.22%
Incidents with Resources Mobilised	413 (0.9%)	703 (1.6%)	+290	+70.22%
Incidents with Resources Arrived at Scene	407 (0.9%)	691 (1.6%)	+284	+69.78%
Total Number of Resources Allocated	1,011 (1.5%)	1,662 (2.6%)	+651	+64.39%
Total Number of Resources Mobilised	966 (1.5%)	1,608 (2.7%)	+642	+66.46%
Total Number of Resources Arrived at Scene	824 (1.6%)	1,341 (2.8%)	+517	+62.74%

In 2016 there were 1,011 resources allocated for purple incidents. In 2017 this had risen by 64.4% to 1,662 resources. This represented 1.5% and 2.6% of all resources allocated in 2016 and 2017, respectively.

In both 2016 and 2017, not all resources allocated were actually mobilised, going from 1,011 resources allocated to 966 being mobilised in 2016 (95.5% of allocated resources being mobilised) and 1,662 allocated but 1,608 mobilised in 2017 (96.7% of allocated resources being mobilised).

There was a reduction in the number of resources that arrived at the scene compared to those that were mobilised (the same pattern was also seen for those allocated versus arrived at scene). In 2016, 824 resources arrived at the scene after 966 were mobilised (85.3% of those mobilised actually arriving). In 2017, 1,341 resources arrived at the scene following 1,608 being mobilised (83.4% of those mobilised actually arriving). The patterns for 2018 were much the same as seen for 2017, although there were increases as noted in the overall incidents and this increase is mirrored in allocations, mobilisation and resources arriving at scene (Table 3).

Table 3 January 2016 and 2018 resource allocation, mobilisation and arrival at scene for purple-coded incidents

Resources	**January 2016 (% versus other codes)**	**January 2018 (% versus other codes)**	**2016 – 2018 Absolute Change**	**2016 – 2018 % Change**
Incidents	414 (0.9%)	916 (1.7%)	+502	+121.26%
Incidents with Resources Allocated	413 (0.9%)	914 (2.0%)	+501	+121.31%
Incidents with Resources Mobilised	413 (0.9%)	914 (2.0%)	+501	+121.31%
Incidents with Resources Arrived at Scene	407 (0.9%)	905 (2.0%)	+498	+122.36%
Total Number of Resources Allocated	1,011 (1.5%)	2,303 (3.6%)	+1,292	+127.79%
Total Number of Resources Mobilised	966 (1.5%)	2,208 (3.6%)	+1,242	+128.57%
Total Number of Resources Arrived at Scene	824 (1.6%)	1,841 (3.6%)	+1,017	+123.42%

In summary, the data indicates that there was an increase in purple-coded incidents in January 2017 and 2018 compared to January 2016.

In terms of resources being allocated, mobilised and arriving on scene there were increases in 2017 and 2018 (just over double the percentage allocation from 1.5% to 3.6% by 2018), although purple calls remain a small percentage of calls and resources allocated. The reasons why there may be differences between resources allocated, mobilised and arriving on scene cannot be assessed with the summary data made available.

Table 4 January 2016 and 2017 resource allocation, mobilisation and arrival at scene for red-coded incidents

Resources	January 2016 (% versus other codes)	January 2017 (% versus other codes)	2016 – 2017 Absolute Change	2016 – 2017 % Change
Incidents	12,847 (26.5%)	4,167 (8.6%)	-8,680	-67.56%
Incidents with Resources Allocated	12,818 (27.9%)	4,139 (9.1%)	-8,679	-67.71%
Incidents with Resources Mobilised	12,805 (28.2%)	4,126 (9.2%)	-8,679	-67.78%
Incidents with Resources Arrived at Scene	12,627 (29.4%)	4,038 (9.5%)	-8,589	-68.02%
Total Number of Resources Allocated	19,952 (29.9%)	7,149 (11.2%)	-12,803	-64.17%
Total Number of Resources Mobilised	19,058 (30.3%)	6,772 (11.4%)	-12,286	-64.47%
Total Number of Resources Arrived at Scene	16,765 (32.2%)	5,703 (12.0%)	-11,062	-65.98%

3.1.3 Red Incidents

The number of red incidents allocated resources in January 2017 was 4,167, a 68% decrease compared to the same period in 2016. This change mirrors the overall incident decrease described above and in Table 1 and is true for resources mobilised and resources arriving at the scene (Table 4, 'Incidents with Resources'). In both 2016 and 2017, there were a small percentage of red incidents that had no resource allocated after being logged as an incident, with this reduction higher in 2017 (0.7% of incidents) compared to 2016 (0.2%). For both 2016 and 2017 there was little difference between resources allocated mobilised and/or arriving on scene (Table 4). This mirrors the results for purple-coded incidents.

When examining the total number of resources that were allocated (where multiple resources may attend each incident), there is evidence of a higher proportion of red incidents not having resources allocated, mobilised or arriving at scene in 2017 compared to 2016. In 2016, 95.5% of red incidents had resources mobilised after initial allocations, but this was 94.7% in 2017. In terms of resources arriving at scene after mobilisation, 88.0% arrived in 2016, but only 84.2% in 2017. The same pattern was present for resources allocated versus arriving. The patterns for 2018 were much the same as seen for 2017, although there were increases as noted in the overall incidents and this increase is mirrored in allocations, mobilisation and resources arriving at scene (Table 5).

In summary, the data indicates that there was a decrease in the number and proportion of red-coded incidents in January 2017 and 2018 compared to January 2016 (as expected in Figure 2). There also appears to have been a reduction in resources arriving after allocation and mobilisation for these red-coded incidents.

Table 5 January 2016 and 2018 resource allocation, mobilisation and arrival at scene for red-coded incidents

Resources	January 2016 (% versus other codes)	January 2018 (% versus other codes)	2016 – 2018 Absolute Change	2016 – 2018 % Change
Incidents	12,847 (26.5%)	5,460 (10.3%)	-7,387	-57.50%
Incidents with Resources Allocated	12,818 (27.9%)	5,438 (11.9%)	-7,380	-57.58%
Incidents with Resources Mobilised	12,805 (28.2%)	5,430 (11.9%)	-7,375	-57.59%
Incidents with Resources Arrived at Scene	12,627 (29.4%)	5,354 (12.1%)	-7,273	-57.60%
Total Number of Resources Allocated	19,952 (29.9%)	8,998 (14.1%)	-10,954	-54.90%
Total Number of Resources Mobilised	19,058 (30.3%)	8,626 (14.1%)	-10,432	-54.74%
Total Number of Resources Arrived at Scene	16,765 (32.2%)	7,394 (14.4%)	-9,371	-55.90%

3.1.4 Amber Incidents

As described above, there was a higher number of amber-coded incidents recorded in 2017 compared to 2016, with 8,930 incidents in 2017 compared to just 1,554 in 2016 (474% increase). This pattern was consistent across allocation, resources mobilised and resources arriving on scene for both years (Table 6).

In terms of the total number of resources per incident, the same pattern was seen with some resources not being mobilised or arriving on scene after allocation. In 2016 4.95% of resources were not mobilised and 12.43% did not arrive on scene after allocation. 15.97% of resources did not arrive on scene after initial allocation. 2017 showed similar patterns (5.27%, 20.37% and 15.79%, respectively), although these resource drop-offs being greater in 2017. The patterns for 2018 were much the same as seen for 2017, although there were increases as noted in the overall incidents and this increase is mirrored in allocations, mobilisation and resources arriving at scene (Table 7). It is possible that resources not arriving at scene had been redirected to a higher priority code while en route, in line with the principles of the NCRM. However, this cannot be assessed with this type of summary data.

Table 6 January 2016 and 2017 resource allocation, mobilisation and arrival at scene for amber-coded incidents.

Resources	January 2016 (% vs other codes)	January 2017 (% vs other codes)	2016 – 2017 Absolute Change	2016 – 2017 % Change
Incidents	1,554 (3.2%)	8,930 (18.4%)	+7,346	+474.65%
Incidents with Resources Allocated	1,552 (3.4%)	8,907 (19.6%)	+7,355	+473.90%
Incidents with Resources Mobilised	1,551 (3.4%)	8,903 (19.8%)	+7,352	474.02%
Incidents with Resources Arrived at Scene	1,547 (3.6%)	8,835 (20.8%)	+7,288	+471.11%
Total Number of Resources Allocated	2,022 (3.0%)	11,917 (18.7%)	+9,895	+489.37%
Total Number of Resources Mobilised	1,922 (3.0%)	11,171 (18.8%)	+9,249	+481.22%
Total Number of Resources Arrived at Scene	1,683 (3.2%)	9,489 (20.0%)	+7,806	+463.81%

In summary, the data indicates that there was an increase in amber-coded incidents in January 2017 and 2018 compared to January 2016. In 2017/18, there also appears to have been a greater reduction in resources arriving after allocation and mobilisation.

Table 7 January 2016 and 2018 resource allocation, mobilisation and arrival at scene for amber-coded incidents

Resources	January 2016 (% vs other codes)	January 2018 (% vs other codes)	2016 – 2018 Absolute Change	2016 – 2018 % Change
Incidents	1,554 (3.2%)	11,623 (22.0%)	+10,069	+647.94%
Incidents with Resources Allocated	1,552 (3.4%)	11,519 (25.1%)	+9,967	+642.20%
Incidents with Resources Mobilised	1,551 (3.4%)	11,510 (25.2%)	+9,959	+642.10%
Incidents with Resources Arrived at Scene	1,547 (3.6%)	11,408 (25.7%)	+9,861	+637.43%
Total Number of Resources Allocated	2,022 (3.0%)	15,586 (24.4%)	+13,564	+670.82%
Total Number of Resources Mobilised	1,922 (3.0%)	14,962 (24.5%)	+13,040	+678.46%
Total Number of Resources Arrived at Scene	1,683 (3.2%)	12,943 (25.3%)	+11,260	+669.04%

Further analysis of individual data was undertaken to ascertain reasons for the difference between the number of calls to SAS in 2016, 2017 and 2018, and the incidents allocated, mobilised and arrived at scene. Table 8 summarises the reasons why incidents were unallocated, not mobilised or did not arrive on scene. However, it was not possible to link individual calls or colour codes with specific reasons given the summary data available for this analysis.

Table 8 The reasons why incidents are unallocated, not mobilised and did not arrive at scene in 2016, 2017 and 2018

	2016	**2017**	**2018**
Passed to NHS24	1,943	1,869	2,058
Not Required Good Intent[2]	1,760	1,639	1,560
Cancelled by Caller[3]	448	621	1,236
Other	415	594	628
Cancelled by Police	251	294	612
CSD Self Care Advice	210	257	186
Downgraded - Emergency to Urgent	118	171	319
CSD Patient Making Own Way	114	201	231
CSD Refer to Alternative Pathway	72	88	348
Malicious Call	28	103	291
Cancelled by Dr/Nurse	37	38	346
N/A	32	29	279
CSD Refer to GP	44	67	133
Patient Made Own Way	35	62	136
Patient Refused Transportation	32	35	33
Patient Not Found	47	16	33
Cancelled by Fire	14	9	20
Advice	<5	7	<5
Not Required- Patient Deceased	6	<5	<5
Treated At Scene	<5	<5	<5
Downgraded - Emergency to Routine	<5	<5	0
Passed to PTS	<5	<5	0
Treat and Refer Procedures used	<5	<5	<5
See and Treat Procedures Used	<5	<5	0
Sum:	5,619	6,110	8,460

[2] 'Not Required Good Intent' refer to incidents where a SAS resource arrived at scene and it was determined that there was no requirement for an ambulance as well.

[3] 'Cancelled by Caller' is where a call has been cancelled prior to a resource arriving scene.

3.2 Are patients with Immediately Life Threatening (ILT) conditions more quickly and accurately identified?

In order to determine if patients with ILT conditions were more quickly and accurately identified, the data were analysed to assess 'speed' and 'accuracy' of identification. 'SAS data' from January 2016 (pre new model adoption) was compared with data from January 2017 and 2018 (post new model adoption) to examine if there were any significant changes in the speed of identifying patients. Data were examined for the purple code, specifically cardiac arrest. The same analyses were conducted for the remaining colour codes: breathing difficulties (red), stroke (amber) and falls (yellow) to ensure any changes pre- and post-NCRM did not have negative impacts on non-ILT conditions. Mean average values were compared over the time points using an independent t-test to assess 'speed' of identification, with assumed equal variances. As these data are normally/approximately-normally distributed, this test is appropriate to perform. Practical significance of the time differences needs to be considered in conjunction with statistical significance i.e. is a mean difference of less than a minute of practical significance for the service and patients.

'Outcomes data' from 2016 and 2017 were analysed to assess 'accuracy' of identification by matching the SAS categorisation with eventual outcome (e.g. non-conveyance, death etc.). Here, a sensitivity/specificity table was created, and measures of accuracy used to compare between the two-time periods.

<u>Statistical Analysis Used: independent t-tests</u>
Independent t-tests ('t' value) are used to compare the means of two independent groups. A p-value of less than 0.05 suggests we can reject the null hypothesis, for example mean arrival times are not significantly different between 2016 and 2017

3.2.1 Category level results – comparison of means ('more quickly identified')

Call started to resource allocation was used to measure 'identification' here. We have also included overall response times to better understand the full picture of the SAS response following identification. In 2018 there were 4,327 more calls than in 2016 (8.9%). Call volume was almost identical between 2016 and 2017, therefore the comparison between these years is possibly a more reliable marker of speed of identification differences pre- and post-NCRM.

There were no statistically significant differences in the timings associated with all stages of dealing with purple-coded calls in January 2016 compared to January 2017, with full results presented in Table 9**.** Mean times for purple calls being started to the first resource arriving at scene were 8.58 minutes (standard Deviation/SD = 4.36 minutes) in 2016 and 8.55 minutes (SD=4.83 minutes) in 2017 (t=0.10, p=0.918). If this is broken down, it can be seen that mean times for the first resource to be allocated after the call was started were 1.48 minutes (SD = 1.38 minutes) in 2016 and 1.66 minutes (SD=2.30 minutes) in 2017 (t=1.45, p=0.149), therefore showing no difference.

The January 2018 data reveals a similar pattern of non-significance, except that a significantly longer period of time in 2018 is taken to allocate the first resource once the call has started. Mean times for purple calls being started to first resource being allocated being 1.48 minutes (SD=1.30 minutes) in 2016, and 2.17 minutes (SD=4.98 minutes) in 2018 (t=2.77, p=0.006) (Table 10).

Red calls in January 2017 took significantly longer for resources to arrive (after the call was started) compared to red calls in January 2016, with a mean difference of 1.27 minutes (95% CI = 1.03; 1.51). In January 2016 the mean time for these resources to arrive on the scene after the call started was 9.12 minutes (SD=5.79), but this had increased to 10.39 minutes (SD=9.06 minutes) by January 2017 (t=10.44, p<0.001). This difference between years appears to be largely driven by the time difference between the call starting and the first resource being allocated (i.e. identification) (t=15.73, p<0.001), as there was no statistically significant difference between 2016 and 2017 for the mean time taken for resources to arrive after being allocated (t=0.49, p=0.621).

Mean resource allocation took 1.65 minutes (SD=1.83 minutes) in 2016 and 2.58 minutes (SD=5.86 minutes) in 2017 after the call was started. The mean time for resources to arrive after allocation was 7.92 minutes (SD=5.24 minutes) in January 2016 and 7.87 minutes in

January 2017 (SD=6.42 minutes). In line with findings from the January 2017 data, red calls in January 2018 took significantly longer for resources to arrive (after the call was started) compared to red calls in January 2016. The mean difference has increased from 1.27 minutes (difference between January 2017 data mean and January 2016 mean) to 4.58 minutes (95% CI = 4.44; 4.98) (Table 9). This difference in years seems to be driven by a statistical increase in timings for all intermediary stages between call starting to resource arriving at scene. Most relevant to this question, call started to first resource allocated mean difference was 3.67 minutes longer in 2018 compared to 2016 and was statistically significant (95% CI=3.95;3.39, $p<0.001$). Therefore, identification of red calls (calls starting to allocation) was slower in 2017 and 2018 compared to 2016.

For amber-coloured calls, there were no statistically significant differences in the timings associated with all stages of resource allocation (and arrival at scene) in January 2016 compared to January 2017. However, for all stages between call started to resource arriving at scene, there was a significant statistical difference between means for January 2018 and January 2016. In all cases, the mean time taken had significantly increased. Specifically, mean difference in call started to first resource allocated had increased from 2.46 to 11.59 minutes (SD=1.38 and 4.98). As noted above, the increase in call volume is a likely factor here.

Category level results summary

Summarising the results just presented, it would appear from the data that changes in the model have not had an impact on the speed with which ILT (purple) calls are identified and responded to. Increases in time taken were seen in 2018, but a likely confounder is the call volume increase in this January. However, it may have impacted on other coded calls with average length of time of allocation (and response) being statistically significantly longer in 2017 and 2018. It is not possible with the summary data to directly link identification and response times to clinical outcomes, which would help to reveal if these changes are having any negative impacts on the outcomes for patients.

Table 9 T-test results for incident responses by colour codes comparing 2016 to 2017

	2016 Mean Time (mins)	2017 Mean Time (mins)	2016 – 2017 Mean Difference (mins)*	95% CI	t	Degrees of Freedom	p-value
Call Started to First Resource Allocated							
Purple	1.48	1.66	-0.18	-0.42; 0.07	1.44	1,113	0.149
Red	1.65	2.58	-0.93	-1.05; -0.81	15.76	16,926	**<0.001**
Amber	2.46	2.77	-0.31	-0.63; 0.01	1.91	10,429	0.057
Yellow	3.50	4.79	-1.29	-1.46; -1.12	15.02	56,257	**<0.001**
Green	5.69	7.07	-1.38	-2.56; -0.20	2.30	5,892	**0.022**
First Resource Allocated to First Response Arrived							
Purple	7.12	6.98	+0.14	-0.40; 0.68	0.51	1,091	0.610
Red	7.92	7.87	+0.05	-0.15; 0.25	0.49	15,940	0.622
Amber	11.74	11.38	+0.36	-0.12; 0.84	1.47	10,371	0.141
Yellow	12.05	13.11	-1.06	-1.25; -0.87	10.71	54,892	**<0.001**
Green	36.75	40.92	-4.17	-7.87; -0.47	2.21	2,180	**0.027**
Call Started to First Resource Arrived on Scene							
Purple	8.58	8.55	+0.03	-0.54; 0.60	0.10	1,095	0.918
Red	9.12	10.39	-1.27	-1.51; -1.03	10.43	16,660	**<0.001**
Amber	14.17	14.13	+0.04	-0.54; 0.62	0.13	10,380	0.893
Yellow	15.47	17.82	-2.35	-2.61; -2.09	17.67	55,004	**<0.001**
Green	48.87	55.82	-6.95	-11.40; -2.50	3.07	2,184	**0.002**
Resource Arrived at Scene to Clear (not conveying to hospital)							
Purple	49.26	51.03	-1.77	-4.59; 1.05	1.23	1,640	0.218
Red	40.14	38.88	+1.26	0.03; 2.49	2.01	8,885	**0.044**
Amber	42.10	45.66	+1.26	-1.27; 3.79	0.98	2,803	0.327
Yellow	40.66	42.57	-1.91	-2.66; -1.16	4.97	20,695	<0.001
Green	33.16	34.55	-1.39	-3.94; 1.16	1.07	1,310	0.284
Resource Arrived at Scene to Leaving Scene (conveying to hospital only)							
Purple	31.01	32.04	-1.03	-3.66; 1.60	0.77	521	0.443
Red	26.65	28.52	-1.87	-2.42; -1.32	6.68	1,2848	**<0.001**
Amber	25.39	29.21	-3.82	-4.55; -3.09	10.27	8,386	**<0.001**
Yellow	25.77	27.70	-1.93	-2.21; -1.65	13.35	40,773	**<0.001**
Green	29.11	31.77	-2.66	-4.75; -0.57	2.50	972	**0.013**
Resource Leaving Scene to Arrival at Destination (conveying to hospital only)							
Purple	14.61	15.29	-0.68	-2.81; 1.45	0.63	521	0.530
Red	18.32	17.74	+0.58	0.01; 1.15	2.01	12,849	**0.044**
Amber	18.30	18.49	+9.19	-0.93; 0.55	0.51	8,384	0.612
Yellow	18.58	19.40	-0.82	-1.08; -0.56	6.21	40,752	**<0.001**
Green	19.25	17.80	+1.45	-0.27; 3.17	1..65	975	0.099
Resource Arrival at Destination to Clear (conveying to hospital only)							
Purple	38.86	37.86	+1.00	-2.46; 4.46	0.57	521	0.57
Red	26.16	29.68	-3.52	-4.10; -2.94	11.81	12,853	**<0.001**
Amber	24.35	27.35	-3.00	-3.80; -2.20	7.34	8,388	**<0.001**
Yellow	24.39	26.82	-2.43	-2.73; -2.13	15.99	40,786	**<0.001**
Green	22.95	25.48	-2.53	-4.28; -0.79	2.85	975	**0.005**

* A positive value indicates 2016 had longer mean times, while a negative value indicates 2017 had longer mean times

Table 10 T-test results for incident responses by colour codes comparing 2016 to 2018

	2016 Mean Time (mins)	2018 Mean Time (mins)	2016 – 2018 Mean Difference (mins)*	95% CI		t	Degrees of Freedom	p-value
Call Started to First Resource Allocated								
Purple	1.48	2.17	-0.69	-1.18;	-0.20	2.77	1,324	**0.006**
Red	1.65	5.32	-3.67	-3.95;	-3.39	25.55	18,230	**<0.001**
Amber	2.46	11.59	-9.13	-10.76;	-7.50	11.03	13,048	**<0.001**
Yellow	3.50	15.71	-12.21	-12.61;	-11.81	59.75	54,822	**<0.001**
Green	5.69	39.83	-34.14	-37.05;	-31.23	23.02	3,807	**<0.001**
First Resource Allocated to First Response Arrived								
Purple	7.12	6.81	0.31	-0.29;	0.88	1.07	1,307	0.285
Red	7.92	8.42	-0.50	-0.70;	-0.30	4.85	17,240	**<0.001**
Amber	11.74	12.62	-0.88	-1.48;	-0.28	2.86	12,944	**0.004**
Yellow	12.05	14.87	-2.82	-3.05;	-2.59	24.07	53,290	**<0.001**
Green	36.75	23.83	12.92	8.70;	17.14	6.01	1,605	**<0.001**
Call Started to First Resource Arrived on Scene								
Purple	8.58	8.98	-0.40	-1.176;	0.38	1.01	1,309	0.311
Red	9.12	13.70	-4.58	-4.94;	-4.22	25.15	1,776	**<0.001**
Amber	14.17	24.14	-9.97	-11.78;	-8.16	10.80	12,952	**<0.001**
Yellow	15.47	30.29	-14.82	-15.30;	-14.34	60.42	53,420	**<0.001**
Green	48.87	76.89	-28.02	-34.84;	-21.20	8.07	1,607	**<0.001**
Resource Arrived at Scene to Clear (not conveying to hospital)								
Purple	49.26	52.38	-3.12	-6.47;	0.23	1.83	1,988	0.068
Red	40.14	41.46	-1.32	-2.42;	-0.22	2.35	9,524	**0.019**
Amber	42.10	50.29	-8.19	-10.85;	-5.53	6.03	4,340	**<0.001**
Yellow	40.66	46.41	-5.75	-6.49;	-5.01	15.25	21,066	**<0.001**
Green	33.16	41.24	-8.08	-11.37;	-4.79	4.82	1,005	**<0.001**
Resource Arrived at Scene to Leaving Scene (conveying to hospital only)								
Purple	31.01	36.69	-5.68	-8.41;	-2.95	4.09	673	**<0.001**
Red	26.65	29.72	-3.07	-3.58;	-2.56	11.85	1,390	**<0.001**
Amber	25.39	30.52	-5.13	-5.93;	-4.33	12.57	10,300	**<0.001**
Yellow	25.77	29.15	-3.38	-3.68;	-3.08	22.38	39,022	**<0.001**
Green	29.11	30.75	-1.64	-4.21;	0.93	1.25	688	0.211
Resource Leaving Scene to Arrival at Destination (conveying to hospital only)								
Purple	14.61	15.46	-0.85	-3.07;	1.37	0.75	669	0.453
Red	18.32	18.51	-0.19	-0.70;	0.31	0.73	13,901	0.463
Amber	18.30	19.19	-0.89	-1.63;	-0.15	2.35	10,298	**0.019**
Yellow	18.58	20.56	-1.98	-2.26;	-1.70	13.94	39,008	**<0.001**
Green	19.25	19.09	0.16	2.42;	2.10	0.14	689	0.89
Resource Arrival at Destination to Clear (conveying to hospital only)								
Purple	38.86	41.00	-2.14	-5.68;	1.40	1.19	672	0.236
Red	26.16	33.17	-7.01	-7.58;	-6.43	24.09	13,908	**<0.001**
Amber	24.35	31.07	-6.72	-7.76;	-5.68	12.61	10,305	**<0.001**
Yellow	24.39	28.54	-4.15	-4.47;	-3.84	25.78	39,029	**<0.001**
Green	22.95	25.30	-2.35	-4.47;	-0.23	2.17	689	**0.030**

* A positive value indicates 2016 had longer mean times, while a negative value indicates 2018 had longer mean times

3.2.2 Condition level results – comparison of means ('more quickly identified')

Data were analysed for four separate conditions, specifically:

- Cardiac arrest (Purple)
- Breathing difficulties (Red)
- Stroke (Amber)
- Falls (Yellow)

Full results are presented in Table 11 for 2016 *vs* 2017 and Table 12 for 2016 *vs* 2018. As above, call started to resource allocation was used to measure 'identification' here. We have also included overall response times to better understand the full picture of the SAS response following identification. Again, the comparison between 2016 and 2017 is possibly a more reliable marker of speed of identification differences pre- and post-NCRM (than 2016 versus 2018) given the call volumes. For the specific conditions examined, there were only two components of the SAS response system where timings associated with the 2017 data were statistically significantly longer than those for the January 2016.

When data from January 2016 were compared with January 2017, there was no difference in means concerning all stages associated with call started to first resource arrived on scene for cardiac arrests, breathing difficulties or stroke (this includes call started to resource allocation aka 'identification'). There was a small, but significant increase in time taken for falls (Table 11). When comparing the January 2018 data with January 2016 there has been a significant increase in timings associated with specific conditions for call started to resource arriving on scene (Table 12).

For breathing difficulties, stroke and falls, the increase in mean value between call started and first resource arrived on scene is statistically significant, with mean increases of 1.4 (95% CI=1.05; 1.87), 3.3 (95% CI=1.65; 5.00), and 16.75 (95% CI=15.50; 18.00) minutes respectively.

For breathing difficulties and falls, the drivers of these increases are both mean time associated with call started to first resource allocated (i.e. identification) [0.99 minutes (95% CI= 0.87; 1.11) and 13.19 minutes (95%CI=12.24; 14.24) respectively] and first resource allocated to first response arrived [0.47 minutes (95% CI=0.08; 0.86) and 3.84 minutes (95% CI=3.31; 4.37) respectively].

For strokes, the driver was the mean time for call started to first resource allocated (identification) with a mean increase of 2.87 minutes (95% CI=2.35; 3.49). For cardiac arrests, the mean increase in time from call started to first resource allocated was statistically significant (0.39 minutes (95% CI=0.22; 0.56). However, this did not have an impact on the overall mean time associated with call started to first resource arrived on scene, with no statistical significance detected in this mean value for January 2018 and January 2016.

Table 11 T-test results for incident responses by conditions comparing 2016 to 2017

	2016 Mean Time (mins)	2017 Mean Time (mins)	Mean Time Difference (2016 - 2017)*	95% CI		t	Degrees of Freedom	p-value
Call Started to First Resource Allocated								
Cardiac Arrest	1.44	1.60	-0.16	-0.39;	0.07	1.36	950	0.176
Breathing Difficulty	1.48	1.64	+0.12	-0.19;	0.41	0.73	2,137	0.470
Stroke	2.43	2.32	+0.11	-0.19;	0.41	0.72	2,137	0.469
Falls	3.61	4.38	-0.67	-1.05;	-0.29	3.49	8,240	**<0.001**
First Resource Allocated to First Response Arrived								
Cardiac Arrest	7.09	7.06	+0.03	-0.54;	0.60	0.10	934	0.918
Breathing Difficulty	8.02	7.75	+0.27	-0.19;	0.73	1.16	3,100	0.245
Stroke	11.18	10.62	+0.56	-1.16;	0.04	1.84	2,119	0.066
Falls	11.44	12.55	-1.11	1.15;	-0.71	5.42	8,067	**<0.001**
Call Started to First Resource Arrived on Scene								
Cardiac Arrest	8.52	8.60	-0.08	-0.69;	0.53	0.26	936	0.796
Breathing Difficulty	9.49	9.33	+0.16	-0.32;	0.64	0.66	3,103	0.509
Stroke	13.58	12.96	+0.62	-0.05;	1.29	1.82	2,120	0.069
Falls	14.97	16.94	-1.97	-2.53;	-1.41	6.86	8,074	**<0.001**

* A positive value indicates 2016 had longer mean times, while a negative value indicates 2017 had longer mean times

Table 12 T-test results for incident responses by conditions comparing 2016 to 2018

	2016 Mean Time (mins)	2018 Mean Time (mins)	Mean Time Difference (2016 - 20178*	95% CI		t	Degrees of Freedom	p-value
Call Started to First Resource Allocated								
Cardiac Arrest	1.44	1.83	-0.39	-0.56;	-0.22	4.47	1,114	**<0.001**
Breathing Difficulty	1.48	2.47	-0.99	-1.11;	-0.87	15.96	3,555	**<0.001**
Stroke	2.43	5.30	-2.87	-3.40;	-2.34	10.73	2,198	**<0.001**
Falls	3.61	16.80	-13.19	14.24;	-12.14	24.59	8,478	**<0.001**
First Resource Allocated to First Response Arrived								
Cardiac Arrest	7.09	6.72	+0.37	-0.20;	0.94	1.27	1,101	0.205
Breathing Difficulty	8.02	8.49	-0.47	-0.85;	-0.09	2.40	3,533	**0.016**
Stroke	11.18	11.58	-0.400	-1.07;	0.21	1.17	2,179	0.241
Falls	11.44	15.28	-3.84	-4.37;	-3.31	14.12	8,303	**<0.001**
Call Started to First Resource Arrived on Scene								
Cardiac Arrest	8.52	8.54	-0.02	-0.62;	0.58	0.07	1,102	0.948
Breathing Difficulty	9.49	10.95	-1.46	-1.87;	-1.05	7.07	3,534	**<0.001**
Stroke	13.58	16.89	-3.31	-5.00;	-1.65	3.93	1,043	**<0.001**
Falls	14.97	31.72	-16.75	-18.00;	-15.50	26.41	8,311	**<0.001**

* A positive value indicates 2016 had longer mean times, while a negative value indicates 2018 had longer mean times

Condition level results summary

The results reveal that the new response model in 2017 maintained average response times for cardiac arrest (and other non-ILT conditions), except for allocations/identification taking longer. However, by 2018 allocation and response times were significantly higher for non-ILT conditions.

3.2.3 Category level results – comparison of count data ('more accurately identified')

Statistical Analysis Used: **Sensitivity and Specificity**

Sensitivity is the ability of a test result or assessment to be positive when the disease or condition is present (a true positive). For example, to correctly identify that the condition was a cardiac arrest.

Specificity is the ability of a test result or assessment to be negative when the disease or condition is not present (a true negative). For example, to correctly identify that the condition was not a cardiac arrest.

Accuracy is the overall likelihood that a patient will be correctly classified.

These results are typically presented in a sensitivity and specificity table:

2016	**Confirmed cardiac arrest**	**Not cardiac arrest**
Allocated a cardiac arrest code	True Positives ('a')	False Positives ('c')
Not allocated a cardiac arrest code	False negatives ('b')	True Negatives ('d')

We can then calculate several measures that indicate the effectiveness of the test or assessment:

Statistic	**Formula**
Sensitivity	$\frac{a}{a+b}$
Specificity	$\frac{d}{c+d}$
Accuracy	$\frac{a+d}{a+b+c+d}$

Sensitivity and specificity tables were created for all purple incidents with outcome data available for January 2016 (Table 13) and January 2017 (Table 14). January 2018 data were not available at the time of analysis. True positives were defined as being coded purple and either the patients died at scene, died in the ED or were sent to an ED and received an onward admission to another ward in the hospital.

False positives were defined as being coded purple, but patients were either not conveyed to hospital or if conveyed, were discharged directly from the ED. False negatives were defined as being coded as something else besides purple (red, amber, yellow, green or unknown) and either the patients died at scene, died in the ED or were sent to the ED and received an onward admission to another ward in the hospital. True negatives were defined as being coded as something else besides purple, but patients were either not conveyed to hospital or if conveyed, were discharged directly from the ED. It should be noted that these definitions are not perfect as there is no direct indication of a patient having a life-threatening condition based on their outcomes data. Dying at scene or in the ED is obviously an effective indicator of a condition being life-threatening and non-conveyance to hospital and direct ED discharge are relatively robust indicators of a non-life-threatening condition. While ED onward admission is an indicator of a need to treat and a level of severity, many non-life-threatening conditions will require ED onward admission. However, without individual-level data is not possible to cross-check SAS coding with outcomes data. This weakness needs to be considered when interpreting the following results. Confidence Intervals (CI) were calculated as 'exact' Clopper-Pearson Confidence Intervals (Erdoğan & Gülhan, 2016).

<u>Statistical Analysis Used: Confidence Intervals</u>

Confidence Intervals (CI) are a range of values which are likely to contain the population parameter of interest (as we can only typically measure a sample of the population). CIs are constructed at a confidence level, such as 95%, as there is a level of uncertainty associated with a sampling method. A 95% CI means that we would expect 95% of the interval estimates we calculate in different samples of the population (for example means) to include the 'true' population parameter (for example, the true mean if we sampled the entire population). CIs are similar to p-values in terms of inferring statistical significance, where in order to reject the null hypothesis we are looking for CIs that do not overlap the values of no difference. For example, if there is no difference between 2016 and 2017, then the values of interest for both years will be the same and/or we will see an overlap in terms of their CIs.

In 2016 there were 157 true positives, 65 false positives, 13,075 false negatives and 16,707 true negatives. Sensitivity, where SAS correctly identify a purple condition, was 1.19% (95% CI = 1.01%; 1.39%). Specificity, where SAS correctly identify a non-purple condition, was 99.61% (95% CI = 99.51%; 99.70%). Both sensitivity and specificity are related to the prevalence of purple/life-threatening conditions in the population, which is low in this sample and therefore we would expect relatively low percentages for sensitivity and higher levels for specificity. In 2017 there were 324 true positives, 134 false positives, 13,896 false negatives and 17,187 true negatives. Sensitivity was 2.28% (95% CI = 2.04%; 2.54 %) and specificity was 99.23% (95% CI = 99.08%; 99.35%). Given that the confidence intervals do not overlap for either measure when comparing 2016 and 2017, we can infer that sensitivity (correctly identifying a purple condition) was higher in 2017 compared to 2016, but specificity (correctly identifying a non-purple condition) was lower in 2017.

We can also measure accuracy using the sensitivity and specificity table. Here, accuracy is defined as the likelihood that a patient will be correctly classified (true positive or true negative). In 2016, the accuracy of the response model was 56.21% (95% CI = 55.64%; 56.77%) compared to 55.52% (95% CI = 54.97%; 56.07%) in 2017. Given the confidence intervals overlap, we do not observe any difference in accuracy in 2017 compared to 2016. This lack of difference is also reflected in Table 15 where there has only been a slight increase in false negatives.

Table 13 Sensitivity and Specificity for purple-coded incidents in January 2016

2016	**Confirmed Purple Condition (death or onward ED admission)**	**Not Purple Condition (non-conveyance or ED discharge)**
Coded as Purple	**TRUE POSITIVE** Died at scene = 120 Died in A&E = 23 ED onward admission = 14 **Total = 157**	**FALSE POSITIVE** Not conveyed = 52 ED discharge = 13 **Total = 65**
Coded as Other (lower risk)	**FALSE NEGATIVE** Died at scene = 198 Died in A&E = 70 ED onward admission = 12,807 **Total = 13,075**	**TRUE NEGATIVE** Not conveyed = 6,549 ED discharge = 10,158 **Total = 16,707**

Table 14 Sensitivity and Specificity for purple-coded incidents in January 2017

2017	**Confirmed Purple Condition (death or onward ED admission)**	**Not Purple Condition (non-conveyance or ED discharge)**
Coded as Purple	**TRUE POSITIVE** Died at scene = 212 Died in A&E = 36 ED onward admission = 76 **Total = 324**	**FALSE POSITIVE** Not conveyed = 98 ED discharge = 36 **Total = 134**
Coded as Other (lower risk)	**FALSE NEGATIVE** Died at scene = 165 Died in A&E = 65 ED onward admission = 13,666 **Total = 13,896**	**TRUE NEGATIVE** Not conveyed = 7,088 ED discharge = 10,099 **Total = 17,187**

Table 15 Proportions of true positives, false negatives, false positives and true negatives

Year	True Positive	False Negative	False Positive	True Negative
2016	0.52%	43.58%	0.22%	55.68%
2017	1.03%	44.06%	0.42%	54.49%
Absolute difference (2017 – 2016)	+0.51%	+0.48%	+0.20%	-1.19%
2016 (purple codes only)	70.72%	-	29.28%	-
2017 (purple codes only)	70.74%	-	29.26%	-
Absolute difference (2017 – 2016)	+0.02%	-	-0.02%	-
2016 (coded as other only)	-	43.90%	-	56.10%
2017 (coded as other only)	-	44.71%	-	55.29%
Absolute difference (2017 – 2016)	-	+0.81%	-	-0.81%

3.2.4 Condition level results – comparison of count data ('more accurately identified')

Sensitivity and specificity tables were also produced for cardiac arrest (purple) incidents. Please note that cardiac arrest incidents have only been compared to the three other main conditions where data was available (and not all incidents as above in Section 3.1.2): breathing difficulty (red), stroke (amber) and falls (yellow).

In 2016 there were 154 true positives, 63 false positives, 2,585 false negatives and 2,823 true negatives (Table 16). Sensitivity, where SAS correctly identify a cardiac arrest, was 5.62% (95% CI = 4.79%; 6.55%). Specificity, where SAS correctly identified breathing difficulties, stroke or a fall, was 97.82% (95% CI = 97.22%; 98.32%).

In 2017 there were 264 true positives, 94 false positives, 2,265 false negatives and 2,353 true negatives (Table 17). Sensitivity was 10.44% (95% CI = 9.27%; 11.70%) and specificity was 96.16% (95% CI = 95.32%; 96.88%). Given that the confidence intervals do not overlap for either measure when comparing 2016 and 2017, we can infer that sensitivity (correctly identifying a cardiac arrest) was higher in 2017 compared to 2016, but specificity (correctly identifying a non-cardiac arrest) was lower in 2017. These results match those of purple conditions described in Section 3.3.

Table 16 Sensitivity and Specificity for cardiac arrest-coded incidents in January 2016

2016	**Confirmed Cardiac Arrest Condition (death or onward ED admission)**	**Not Cardiac Arrest (non-conveyance or ED discharge)**
Coded as Cardiac Arrest	**TRUE POSITIVE** Died at scene = 119 Died in A&E = 23 ED onward admission = 12 **Total = 154**	**FALSE POSITIVE** Not conveyed = 50 ED discharge = 13 **Total = 63**
Coded as Other (lower risk)	**FALSE NEGATIVE** Died at scene = 17 Died in A&E = 18 ED onward admission = 2,550 **Total = 2,585**	**TRUE NEGATIVE** Not conveyed = 1,102 ED discharge = 1,721 **Total =2,823**

Table 17 Sensitivity and Specificity for cardiac arrest-coded incidents in January 2017

2017	**Confirmed Cardiac Arrest Condition (death or onward ED admission)**	**Not Cardiac Arrest (non-conveyance or ED discharge)**
Coded as Cardiac Arrest	**TRUE POSITIVE** Died at scene = 199 Died in A&E = 25 ED onward admission = 40 **Total = 264**	**FALSE POSITIVE** Not conveyed = 78 ED discharge = 16 **Total = 94**
Coded as Other (lower risk)	**FALSE NEGATIVE** Died at scene = 8 Died in A&E = 13 ED onward admission = 2,244 **Total = 2,265**	**TRUE NEGATIVE** Not conveyed = 965 ED discharge = 1,388 **Total = 2,353**

Accuracy using the sensitivity and specificity table can also be measured. Here, accuracy is defined at the likelihood that a patient will be correctly classified (true positive or true negative). In 2016, the accuracy of the response model was 52.92% (95% CI = 51.61%; 54.24%) compared to 52.59% (95% CI = 51.19%; 53.99%) in 2017. Given the confidence intervals overlap, we do not observe any difference in accuracy in 2017 compared to 2016. This lack of difference is also reflected in Table 18 where there has only been a slight increase in false negatives (1.25%), but a larger increase in true positives (2.78%).

Table 18 Proportions of true positives, false negatives, false positives and true negatives

Year	True Positive	False Negative	False Positive	True Negative
2016	2.74%	45.96%	1.12%	50.19%
2017	5.31%	45.52%	1.89%	47.29%
Absolute difference (2017 – 2016)	+2.57%	-0.44%	+0.77%	-2.90%
2016 (cardiac arrest codes only)	70.97%	-	29.03%	-
2017 (cardiac arrest codes only)	73.74%	-	26.26%	-
Absolute difference (2017 – 2016)	+2.78%	-	-2.78%	-
2016 (coded as other only)	-	47.80%	-	52.20%
2017 (coded as other only)	-	49.05%	-	50.95%
Absolute difference (2017 – 2016)	-	+1.25%	-	-1.25%

3.3 Are more lives saved as a consequence of the best available resources being dispatched to the patient?

3.3.1 Category level results

Statistical Analysis Used: Kaplan-Meier Survival Probabilities
The Kaplan-Meier survival probability is the probability of surviving in a given length of time while considering time in a series of intervals. For example, if everyone in a population survives, then the probability of survival will be 1 (100%). If everyone dies, the survival probability is 0 (0%). Kaplan-Meier survival plots/curves are typically used to visualise these probabilities.

For this analysis, given that we only had summary data, we do make a significant assumption that 'best available resources' have in fact been despatched and arrived on scene and best available resources link to the changes between 2016, 2017 and 2018. Kaplan-Meier estimates of survival probabilities were computed as a function of time (in hours, where 0 is 'died at the scene', and then we have death within 24, 48, and 720 hours, i.e., 30 days) by category (purple, red, amber, yellow, green and not categorised) and year (2016, 2017 or 2018). The aim was to assess differences in survival ('lives saved') within and across years. The number of deaths after 30 days, as a proportion of calls, are given in Table 19.

Table 19 Number of deaths for each colour code (and as a proportion of all calls within that colour code) 30 days after the 999 call

	2016			2017			2018		
	Deaths	Total Calls	% Deaths	Deaths	Total Calls	% Deaths	Deaths	Total Calls	% Deaths
Purple	194	226	85.8%	335	469	71.4%	502	684	73.4%
Red	647	9,267	7.0%	208	2,717	7.7%	378	3878	9.7%
Amber	88	1,247	7.1%	354	6,916	5.1%	362	7232	5.0%
Yellow	963	20,457	4.7%	1096	21,893	5.0%	1,104	17,442	6.3%
Green	7	205	3.4%	6	176	3.4%	10	472	2.1%
Not Categorised	100	1,462	6.8%	158	2,405	6.6%	597	8,193	7.3%
Cardiac Arrest	191	221	86.4%	299	362	82.6%	339	406	83.5%
Breathing Difficulty	202	2,018	10.0%	79	451	17.5%	220	2,044	10.8%
Stroke	54	730	7.4%	80	1,004	8.0%	92	980	9.4%
Falls	92	2,926	3.1%	152	3,334	4.6%	135	2,067	6.5%

The following plot (Figure 3) describes the probability of survival by category in (January) 2016, 2017 and 2018. In all years, survival for purple-coded patients is markedly lower with respect to all other causes (as one would expect) and reflects that purple-coded calls/conditions are a unique category (in terms of risk of death). The risk of death across the other codes is small in comparison and therefore differences of practical/clinical significance seem to exist only for the purple-coded patients.

The survival probabilities within each category have also been compared between years (Figure 4). Please note that the vertical (y) axis for the purple category shows the full survival probability (0-1), but for the other codes this is condensed to only reflect survival probabilities between 0.9 and 1. Figure 4 shows that there seems to be a considerable (~20%) increase in survival for purple-coded patients from 2016 to 2017 which is constant over time from time 0 (confirmed dead when the ambulance arrives at the scene) to 30-days. When comparing January 2016 to January 2018 for the same group, survival also increased (~10%). The number of lives saved, 30 days post-call, in patients with ILT conditions in January 2016 (pre-NCRM) was 32 (14.2% of purple calls), and in post-NCRM in January 2017 was 134 (28.6% of purple calls) and in January 2018 was 182 (26.6% of purple calls) (Table 19).

Although the numbers of patients with ILT conditions has increased, the data from the specificity and sensitivity analysis (Table 15) shows that there is no difference in false positive rates between the years. This suggests that the acuity of these patients remains very high and that the increase in volume represents patients correctly identified with the highest requirement for immediate response. Therefore, the increase in survival probability with those with ILT conditions is not likely to be caused by artificial inflation caused by conservative allocation of patients with ILT conditions to the purple code. For the remaining colour codes there were no clear differences in survival in 2017/18 *vs* 2016, given the maximum difference was around just 2% (Figure 4, where the scales for these codes reflect survival probabilities between 0.9 and 1)

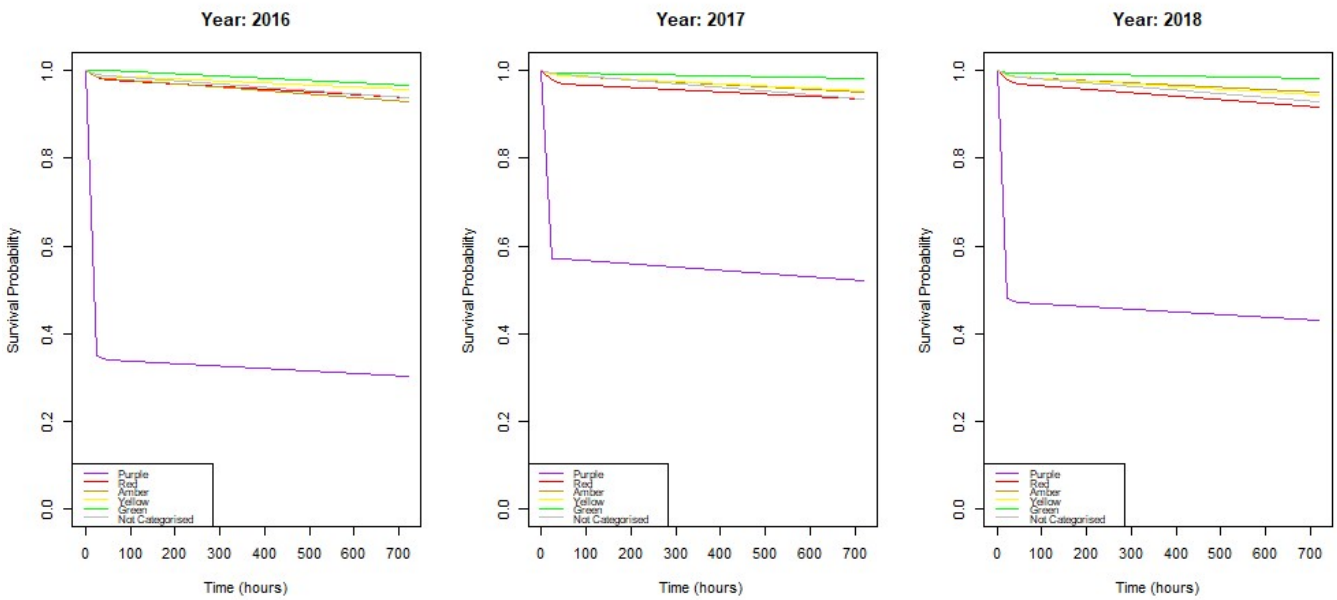


Figure 3 Survival probability plots for all colour codes in 2016, 2017 and 2018

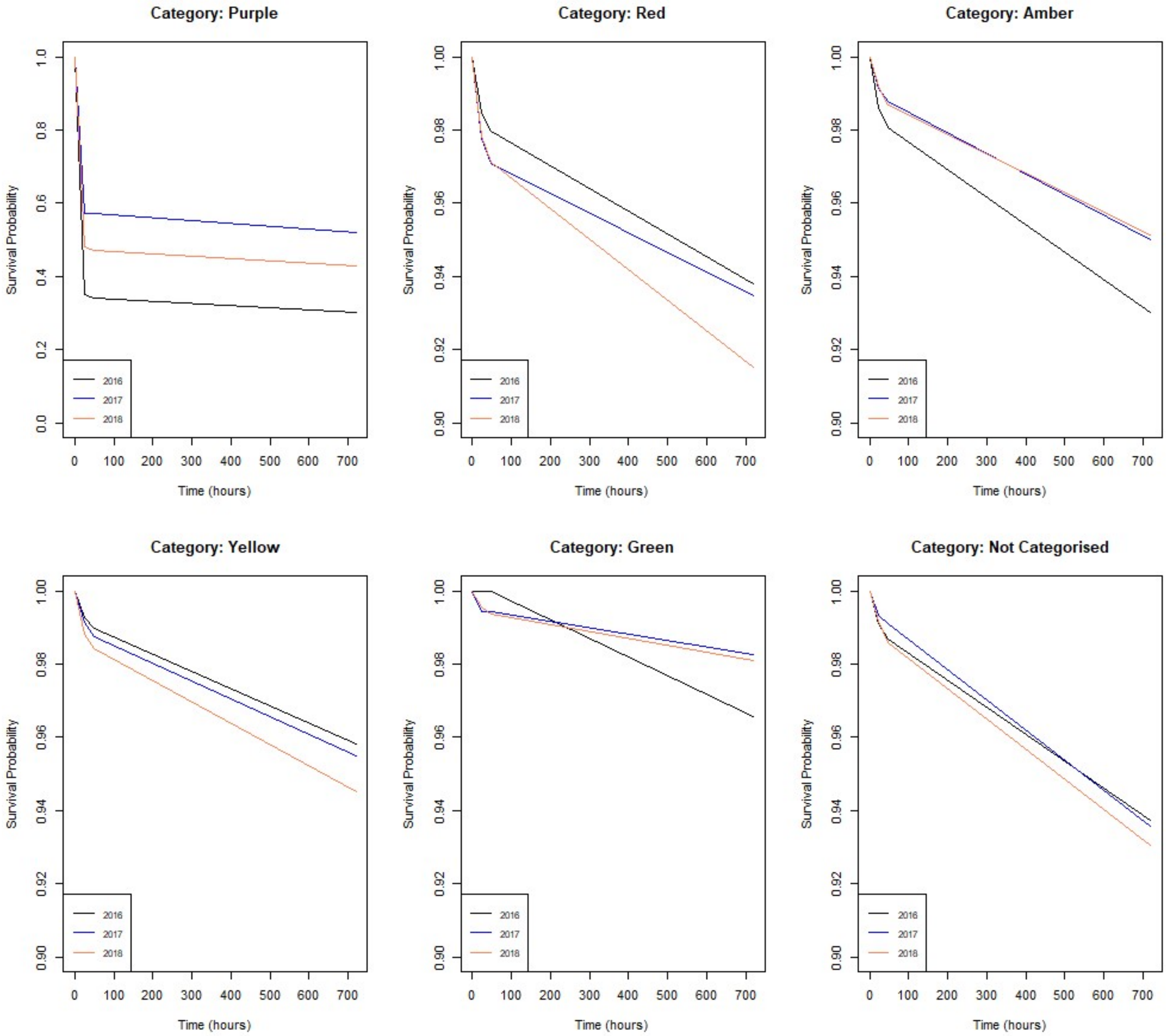


Figure 4 Survival probability plots comparing 2016, 2017 and 2018 for each code category

3.3.2 Condition level results

The same analyses as in Section 3.3.1 were performed for condition-specific incidents: cardiac arrest, breathing difficulties, stroke and falls. The following plots (Figure 5) describe the probability of survival by condition in (January) 2016, 2017 and 2018. In all years, survival for cardiac arrest patients appears markedly lower with respect to all other causes (as one would expect), matching the result for purple codes. In 2017, there appears to be a lower survival from breathing difficulties compared to 2016, which is considered below (Figure 6).

Note that the vertical (y) axis for the cardiac arrest conditions shows the full survival probability (0-1), but for the other codes this is condensed to only reflect survival probabilities between 0.7 and 1 (Breathing Difficulty) and 0.9 and 1 (Stroke and Falls). From this analysis, there seems to be a ~10% increase in survival for cardiac arrest patients from 2016 to 2017 which is constant over time from time 0 (confirmed dead when the ambulance arrives at the scene). However, by 2018, survival has returned to 2016 levels. As with the analysis on purple codes, this increase in survival in 2017 is unlikely to be the result of an artificial inflation caused by conservative allocation of cardiac arrest codes given the decrease in false positives (Table 18).

Breathing difficulty (a sub-set of the red calls) seems to have worsened between 2016 and 2017, with 451 patients having a decrease in survival from 3% to 6%, with the gap widening as time passes. However, by 2018, survival was at 2016 levels despite the number of incidents (n=2044) back to the levels seen in 2016 (n=2018). No remarkable differences seem to be present for stroke and falls (<1% at any point in 2017 and 2% in 2018) (Figure 4).

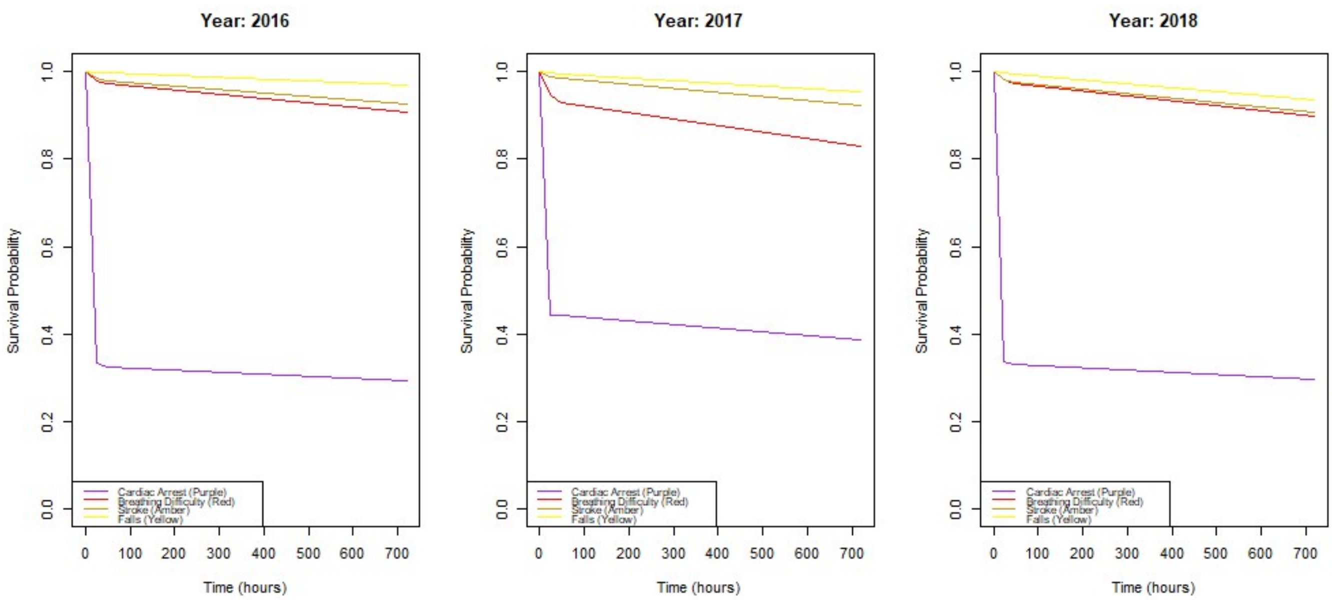


Figure 5 Survival probability plots for conditions in 2016, 2017 and 2018

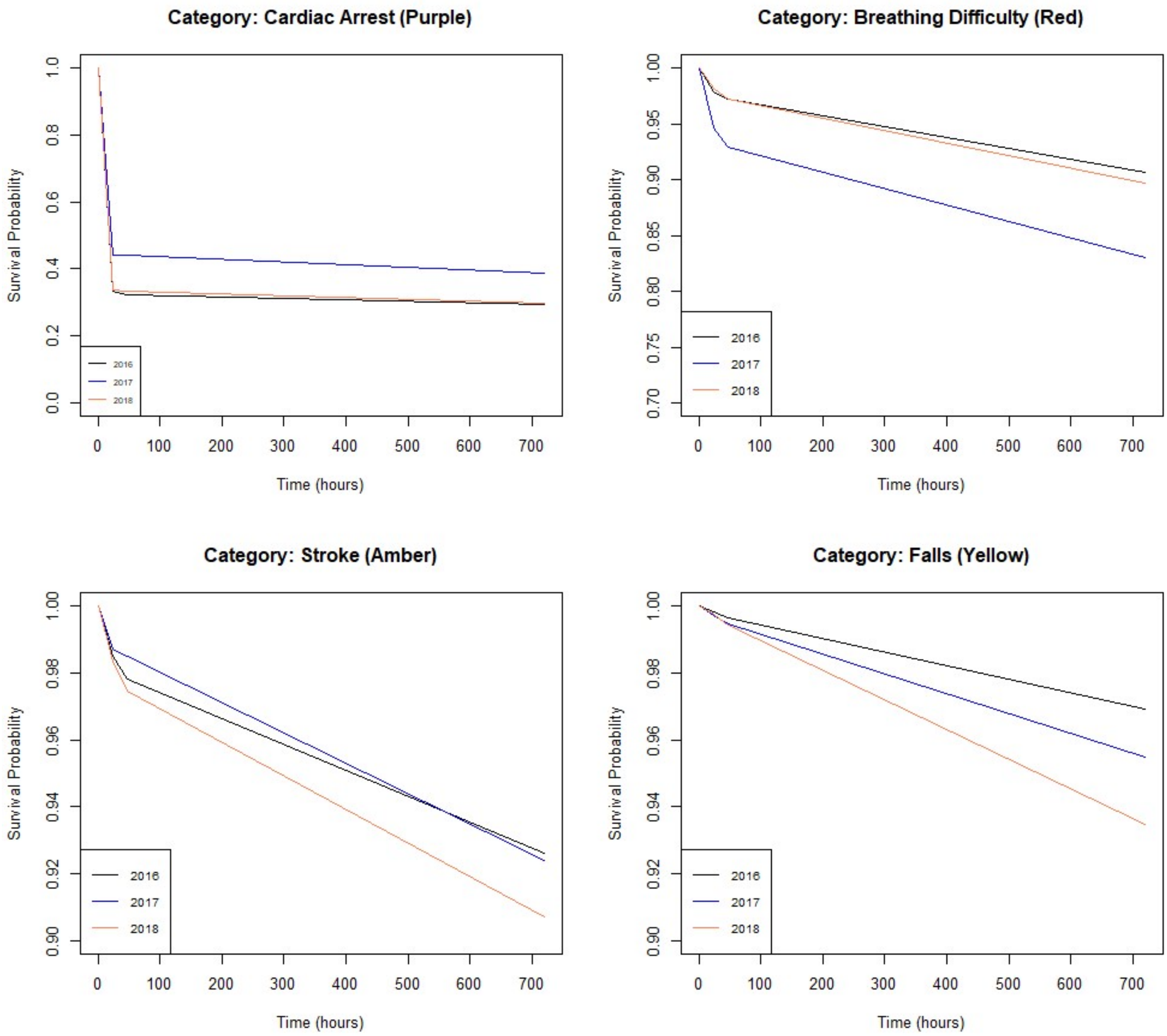


Figure 6 Survival probability plots comparing 2016, 2017 and 2018 for each condition

3.4 Are improved clinical outcomes achieved if the matched resources are sent first time for patients with non-ILT conditions?

Given the summary data available from SAS it is not possible to estimate if resources are matched to need for patients in any category. Furthermore, we do not have data on if these resources were sent first time. The UCDM, ED and NRS data made available for this analysis also do not contain distinct clinical outcomes for patients. While we do have information on admissions and discharge by speciality and ICD-10 conditions, this tells us little about the outcomes of patients beyond mortality within 24, 48 and 720 hours.

The analysis in Section 3.3 covers this, identifying that there has been little or no change in mortality between 2016 and 2017, except for a small improvement for amber incidents (Figure 4). If the assumption is made that in 2017 matched resources are sent first time to patients with non-immediately-life-threatening conditions, then in terms of mortality at least, there does not appear to be any improvement compared to the old response model. For non-ILT conditions there would be little expectation that mortality would improve as it is already very low for these types of calls. It is also reassuring that mortality has not risen in these groups though, and that for amber incidents survival has actually improved, despite SAS taking (intentionally) longer to get to many of these patients following the NCRM. It is not possible to ascertain if there has been an impact on other clinical outcomes besides mortality with these data.

Chapter 4: Key Findings

Context Results

There were several key points relating to the context of the findings:

- There were no differences in the overall number of incidents recorded between January 2016 and January 2017, although total incidents had increased by 8.9% in January 2018 particularly in the purple and red categories. Those with ILT conditions consistently receive care quicker than those patients with lower acuity needs (i.e. not life-threatening conditions).
- The number of purple, amber, yellow and green codes were all higher in 2017 compared to 2016. In contrast, red incidents were lower in 2017. This pattern continues in 2018. This was broadly as intended within the NCRM modelling (Figure 2).
- In 2016 there were 1,011 resources allocated for purple incidents. In 2017 this had risen by 64.4% to 1,662 resources. This represented 1.5% and 2.6% of all resources allocated in 2016 and 2017, respectively. By 2018, this rise continued, with resources allocated at 3.6%, a 128% increase compared to 2016.
- The number of red incidents allocated resources in January 2017 was 4,167, a 68% decrease compared to the same period in 2016. In 2018, this decrease was not as considerable compared to 2016, showing a 57% drop. This reflects the phasing of NCRM with the changes in coding applying in year 1 with no subsequent change into year 2.

Interpreting this data

It should be noted that data is taken from only three (and in some cases two) time points and only from the month of January. While this does allow some relevant comparisons between the years, the findings cannot be generalised to the whole year and the whole time-period in question (January 2016 – January 2018). In addition, call volume was approximately 9% higher in 2018 compared to 2017 and 2016 (which were similar) with over 4,000 more calls in January 2018.

Further analysis of the data using data from each month, as well as individual-level data (rather than it being aggregated), would allow much more robust and relevant evidence of change and the impact on the service and patients.

Are patients with Immediately Life Threatening (ILT) conditions more quickly and accurately identified?

Patients with ILT conditions (purple calls) would appear to be more accurately identified post-NCRM with a noticeable increase in patients coded with ILT conditions by 2018. The time to respond to ILT conditions was slightly longer (but not statistically significant).

Speed

Resource allocation was used as an indicator of speed of identification. We found that resource allocation (and in turn response times) did not differ significantly between January 2016 (pre-NCRM) and January 2017 (post-NCRM introduction) for ILT (purple) calls. However, there was a longer time to allocate resources (i.e. identify) purple calls in 2018 compared to 2016 and this was statistically significant, although as previously stated there was a significant increase in overall and specifically in purple code demand in 2018. For all other codes, 2018 and 2017 resource allocation were also significantly slower than 2016 (except amber 2017 calls) as expected with a priority-based system. Call handlers were provided with further training and development in the process of triage over the course of 2016 onwards, with the aim of more accurately identifying patients into the most appropriate category, and therefore it was to be expected that time to allocate resources and identification into the correct category would take longer.

Accuracy

Comparing 2016 (pre-NCRM) and 2017 (post-NCRM introduction) outcomes data, we found that sensitivity (correctly identifying a purple, ILT condition) was higher in 2017 compared to 2016, but specificity (correctly identifying a non-ILT condition) was lower in 2017. Overall accuracy (the likelihood of being correctly identified as either ILT or non-ILT) was not different between the two-time points. Similar results were also seen for the cardiac arrest cases within the purple calls.

Are more lives saved as a consequence of the best available resources being dispatched to the patient?

Survival for purple-coded patients is markedly lower with respect to all other causes (as one would expect) and reflects that purple-coded calls/conditions are a unique category (in terms of risk of death) and represent the majority of incidents where patients face an immediate threat to life (ILT). The risk of death across the other colour codes is small in comparison and therefore differences of survival seem to exist only for the purple-coded patients.

The cardiac arrest rate within the purple coded is around 53%. Survival analysis for all patients within the purple code and specifically for those affected by cardiac arrest are considered next.

There seems to be a considerable (~20%) increase in survival for all purple-coded patients comparing January 2016 to January 2017, which is constant over time from time 0 (confirmed dead when the ambulance arrives at the scene) to 30 days post-call. When comparing January 2016 to January 2018 for the same group, survival also increased (~10%).

The number of lives saved, 30 days post-call, in patients with ILT conditions in January 2016 (pre-NCRM) was 32 (14.2% of purple calls), and in post-NCRM in January 2017 was 134 (28.6% of purple calls) and in January 2018 was 182 (26.6% of purple calls).

It is possible that the increase in accuracy from the NCRM for purple ILT calls identified in question 1 is contributing to this improvement, although the change in accuracy is quite small. It is also possible that other (unmeasured) factors not fully explored here may also be contributing to this improvement, such as resource allocation and first response treatments. Given the lack of change in response times seen in question 1, response times are unlikely to be contributing. Although the numbers of patients with ILT conditions has increased, the data from the specificity and sensitivity analysis (Table 15) shows that there is no difference in false positive rates between the years. This suggests that the acuity of these patients remains very high and that the increase in volume represents patients correctly identified with the highest requirement for immediate response. Therefore, the increase in survival probability with those with ILT conditions is not likely to be caused by artificial inflation caused by conservative allocation of patients with ILT conditions to the purple code but rather by appropriate allocation and intervention(s) to those patients at risk from death due to ILT conditions. In terms of the 2018 survival probability being lower than in 2017, it is possible that the higher call-load in 2018 has limited the impact previously seen in 2018. Continued monitoring of these data is needed to identify how survival has been impacted by the NCRM over the longer-term.

Are improved clinical outcomes achieved if the matched resources are sent first time for patients with non-ILT conditions?

Overall, survival for all non-ILT codes is similar, as noted above (where purple calls carry much higher risk of death). Between years there appears to be no difference in survival for these remaining codes, with the largest differences being just 2%.

Breathing difficulty (a sub-set of the red calls) seems to have worsened between 2016 and 2017, with a 3% to 6% decrease in survival, with the gap widening as time passes. However, by 2018, survival was back at 2016 levels. No remarkable differences seem to be present for stroke and falls.

Data on further clinical outcomes were not available within this dataset to analyse in any further detail.

Chapter 5: Discussion

Comparison to Other Studies

Other UK Ambulance services have been developing and testing new response models with the same ambitions as SAS to replace performance frameworks with outdated targets with a view to alignment with contemporary pre-hospital and in-hospital health models of care. The Lord Carter of Coles' Review (2018) highlights variation in call handling time and dispatch times, within a context of attracting and retaining control centre workforce in England: therefore, the importance of further research on the 'whole system' is required to further our understanding of producing services that can respond to acuity of illness, the needs of patients, and ultimately produce the best service for the best outcomes for patients.

The introduction of the new response model in Wales was developed following a review of the existing system (McClelland, 2013), a one-year pilot followed by whole country implementation and evaluated in 2017 (PACEC, 2017). Evaluation of a new response model in pilot sites in England was informed by a rapid review of the literature (Turner et al, 2015) and evaluated in 2017 (Turner et al, 2017). The Welsh and English evaluations show improvements in meeting targets related to improved clinical outcomes and more effective and efficient care. Both identified some challenges in structure, process and outcome which are being used to continue development.

Those evaluations were more substantial with wider remits than the evaluation reported here which is confined to routinely collected data, although it is strengthened by linked data to mortality records. Internationally, Ambulance Victoria implemented a revised response model in 2015 following a review of ambulance dispatch in 2015. Victoria Ambulance's three areas of improvement were broadly the same as those of Wales, England and Scotland apart from their inclusion of the availability of specialist care and links to community care. The evaluation in 2017 was substantially positive (Ambulance Victoria, 2017).

Strengths and Weaknesses of the evaluation

The SAS database, in combination with data linkage to ISD records, is unique within the UK making it possible for SAS to conduct an exceptional review of evidence to support the development of their NCRM. The database also made it possible to conduct this evaluation, although with summary data only. Without individual-level data we cannot track the patient journey to see a fully-formed picture

In addition, the data linked between SAS and ISD ('ISD outcomes data') uses the last call on the system to link patients. However, the 'SAS data' contains all registered calls to the system. This means that the number for incidents does not match the two datasets and when analysing the 'SAS data' we are dealing with some incidents with multiple calls. Therefore, some incidents will essentially be double-counted and this will reduce the variation in the dataset (for example three incidents all with the same response time). It should be noted that this has the impact of reducing the standard deviation around the mean, potentially making the estimates here more conservative. Further data checking and sub-group analyses is required to confirm if this is prevalent in all colour codes/conditions and if it alters the conclusions that can be drawn from Section 3.2.

This evaluation has been limited by using data from the month of January across three years. With additional resources and time, the analysis could be improved by analysing data over several months and preferably for every month of the year in a time-series. In addition, as just described, this would ideally be individual level data that could be aggregated when needed for visualising results.

In the sensitivity and specificity analysis (Section 3.2), there was no direct indicator of a patient having a life-threatening condition based on the available outcomes data. Dying at scene/ED is obviously an effective indicator of a condition being life-threatening and non-conveyance to hospital and direct ED discharge are relatively robust indicators of a non-life-threatening condition. While ED onward admission is an indicator of a need to treat and a level of severity, many non-life-threatening conditions will require ED onward admission. However, without individual-level data it was not possible to cross-check SAS coding with outcomes data, meaning this group (onward admission) could not be divided into those that might match purple (or non-purple) coded incidents. This means that categorisation into true positives and false negatives in particular may not be accurate for some cases and is a potential bias in the results and their interpretation.

Interpretation of Results

For this evaluation, January 2016 and January 2017 provided a good comparison as there were similar numbers of calls overall between the two years, supplemented by data from January 2018. There were definite changes between 2016 and 2017/18. However, it is not possible to tell about any re-categorisation. It was the intention of SAS to change the proportion of 999 calls requiring an eight-minute response from 30.6% of the total volume to reflect the acuity of the patient. The modelling meant that calls with a cardiac arrest rate over 10 would be coded purple with the nearest two resources dispatched, and calls with a cardiac arrest rate over 1% coded red with the nearest resource dispatched. In the event purple codes have a cardiac arrest rate over 50% and red codes have a cardiac arrest rate of 1.3%. This resulted in purple/red in 2016: 27.4%; purple/red in 2017: 10.1%; purple/red in 2018: 12.0%.

The modelling, based on a clinical review of ~500,000 cases, resulted in significant changes to the way codes were categorised in terms of response. A number of red and yellow codes moved into the purple category. A significant number of red codes moved to the amber category. Since the launch of the pilot in November 2016 there have been very few subsequent code changes resulting in more stability in the system offering the opportunity for more research to drive improvement.

The analysis for the three January periods appears to show that there has been an improvement in survival for patients categorised purple who are the highest group with risk of death, and stable survival rates for those patients moved from red to amber categories (as would be expected given the lower acuity of patients' needs). While this is positive, we have also seen that the improvements were largest in 2017 and this was not quite maintained into 2018. It is possible that the increased call-load in 2018 has had a knock-on effect on survival in the purple call group, although data would need be collected and analysed over a longer time-period to assess which of the survival patterns (2017 or 2018) are more reliable and consistent over time.

Following the introduction of the new response model, survival for patients on purple/cardiac arrest calls has improved, and these findings suggest that response times by themselves do not account for this improvement. There was a small increase in the number of true positives in 2017 (purple calls that were indeed purple calls) which may explain some of the survival advantage whereby the new response model is more accurate at identifying life threatening conditions. Purple and red calls still have the historical eight-minute target (hence the use of mean data within this research report) however a move to reporting median and 90th percentile response times may allow for comparison (for example to other Ambulance Trusts in England) and more meaningful data in that percentile reporting reduces the effect of extremely large or small values, and so it may give a better idea of a 'typical' value.

Limitations with our accuracy, sensitivity and specificity measures need to be taken into account here given that we analysed aggregated data: indeed, individual-level data that can track patient progress through the system is needed to measure these factors more accurately. Other contributing factors may be the number and resource types being allocated, the treatment given at scene or factors external to SAS, although this was not measurable within this evaluation but could be ideas for further research and evaluation.

Preliminary analysis of the number and type of resources sent to incidents suggest that this is also not a factor. In 2016, 4.07 crew members per call were allocated to purple calls compared to 3.96 in 2017, 3.32 crew/call arriving at scene in 2016 compared to 3.20 in 2017. In terms of specific resource types, 1.47 ambulances per call arrived on scene in 2016 compared to 1.37 in 2017. Paramedic response vehicles were more prevalent in 2017 compared to 2016, with 0.42 arriving per call in 2017 compared to 0.33 in 2016.

Ambulance teams have achieved 39% Return of Spontaneous Circulation (ROSC) events in cardiac arrest patients in 15/16, an increase of 5% from the previous year (Scottish

Ambulance Service, 2016a). This improvement has followed the introduction of a National Strategy in 2015 to improve survival following Out of Hospital Cardiac Arrest. To implement this strategy, SAS has developed a comprehensive work program looking at enhancing all elements of the 'Chain of Survival'. This includes elements both within NCRM such as code categorisation within the model (i.e. purple coding expansion), improved PAD mapping, Pre Entry Question introduction and triple responding of resources, and unrelated elements such as additional staff training and trials of co-responding with Scottish Fire and Rescue Service and Police Scotland, and the continued development of the volunteer Community First Responder program.

Although not the focus of our research questions, we identified that not all resources allocated actually arrive at scene – for the vast majority of these, the allocation wsa deemed to be no longer needed. We undertook further analysis of the data to explain the reasons why this might be so, detailed in Table 8, and one particular interesting point is that there were relatively few cases (n=12 over the 3 years in Table 8) where the ambulance crew were able to 'see and treat' or 'treat and refer'. This may suggest that those calling for ambulance services are in need of conveyance to hospital but may also be due to the lack of usage of this service by crews: the data we had were not able to capture this in more detail but might be an area worthy of further research and exploration.

Recommendations for Further Research

This research provides an overview of the NCRM in place and its effect on response times, and survival rates for a small number of conditions. What this research does not do, is help us understand what makes the system work, in what circumstances it works best, and how it can be further developed, through an iterative process, to provide the best care for patients. The requirement to provide ongoing research and development into this whole system is strongly suggested.

There are several performance framework related areas that are suggested for further research and evaluation of the NCRM:

1. Analysis of time-trends over full years (rather than one month) pre- and post-NCRM
2. An examination of 'whole pathway' patient outcomes using individual-level data across entire years pre- and post-NCRM (e.g. length of hospital stay; readmission rates to hospital)
3. An evaluation of patient experience within each colour category
4. An evaluation of staff experience of the system, including call handling (perhaps where this process could be expedited), ambulance personnel, emergency department staff (and others): this should focus on understanding what makes the system work (or not), and in what circumstances
5. Further exploration of the coding and subsequent care of patients linked to short and longer-term outcome data (especially those data that are routinely collected and could be data linked)
6. Examination of the impact on resources including: shift over-runs; costs of overtime; use of vehicles, shift patterns at time of most need 24/7)
7. Further data analysis of other conditions within each colour category
8. Further comparative analysis of these results with literature on other systems in use internationally
9. Analysis of morbidity data and linked to other clinical outcomes
10. The possible use of other 'see and treat/refer' pathways (e.g. to GP and other community health and social care services) for those with lower acuity conditions and comparison of health and social care outcomes compared to the current use of the system

With the recent Carter Review (2018) of ambulance services in England suggesting there are areas of good practice while a number of variations in quality and delivery of the service, the need for further research and evaluation is strongly advocated. A longitudinal comparative research project, sampling from across service in Scotland and the rest of the UK, is achievable and should consider measuring, recording and data linking clinical indicators and health and social care outcomes to provide information on the best form of care delivery. Large scale longitudinal research that provides a comparison of organisation systems and structures and understanding how these are designed to deliver a whole system approach to service delivery. Findings and potential impact of this research, within the environment of increased demand, could ultimately reduce variation in service and improve health and social care outcomes for patients.

Conclusions

By January 2018 the number of incidents (n=52,871) had increased by 9% when compared to January 2016 (n=48,544), amounting to over 4000 more incidents in 2018 than seen in 2016 or 2017. During this time of high demand in 2017 and particularly 2018, the NCRM does accurately identify patients who have the greatest need for services from SAS. The NCRM's identification and triage of patients into triage categories, although taking time for the call handler and dispatching system, can get the ambulance and its crew to patients with the greatest need and this has improved the survival of those with immediate life-threatening conditions. Those with lower acuity needs are responded to but in a longer time period as expected when using a priority-based system (but with no apparent impact on survival). These conclusions are reached in the context of analysing aggregated data over three fairly short time-periods and further research over a longer time frame, with longitudinal data on individual cases, would further improve the evidence base for the NCRM.

References

Ambulance Victoria (2017) *Revised clinical response evaluation report: delivering our patients the right care, at the right place, at the right time* Department of Research and Evaluation, Ambulance Victoria, Australia

Bigham BL; Kennedy SM; Drennan I; Morrison LJ (2013) Expanding paramedic scope of practice in the community: a systematic review of the literature
Prehospital Emergency Care **17** (36) 361 - 372

Blackwell T; Kilne J; Willis J; Monroe Hicks G (2009) Lack of association between pre-hospital response times and patient outcomes *Prehospital Emergency Care* ***13*** *(4) 444 – 450*

Carter, Lord Carter of Coles *(2018) Operational Productivity and Performance in English NHS Ambulance Trusts: Unwarranted variations.*

Committee on the Treatment of Cardiac Arrest: Current Status and Future Directions; Board on Health Sciences Policy; Institute of Medicine; Graham R, McCoy MA, Schultz AM (editors). *Strategies to Improve Cardiac Arrest Survival: A Time to Act*. Washington (DC): National Academies Press (US); 2015 Sep 29. 1, Introduction. Available from: https://www.ncbi.nlm.nih.gov/books/NBK321509/

Erdoğan S & Gülhan OT (2016) Alternative confidence methods used in diagnostic accuracy studies *Computational and mathematical methods in medicine.* doi: 10.1155/2016/7141050

O'Neill S; Walters A; Steventon A (2017) *The impact of redesigning urgent and emergency care in Northumberland.* The Health Foundation

PACEC with Turner J (2017) *Clinical model pilot evaluation* (final report)
NHS Wales Ambulance Service

McClelland S (2013) *A Strategic Review of Welsh Ambulance Services*
The Welsh Ambulance Service

National Ambulance Resilience Unit (posted 22 November 2016) Scottish Ambulance commences trial of new response system https://naru.org.uk/scottish-ambulance-commences-trial-new-response-system/

Scottish Ambulance Service (2015) *Towards 2020: Taking Care to the Patient*
www.scottishambulance.com/UserFiles/file/TheService/Publications/Strategic%20Plan_Online%20pdf.pdf

Scottish Ambulance Service (2016a) *Scottish Ambulance Service 2016-17 Local Delivery Plan*
http://www.scottishambulance.com/UserFiles/file/TheService/Publications/2016-07-27SASLDP201617.pdf
http://www.scottishambulance.com/UserFiles/file/TheService/Publications/2016-07-27SASLDP201617.pdf

Scottish Ambulance Service (2016b) *Scottish Ambulance Service Annual Report 2015-16* http://www.scottishambulance.com/UserFiles/file/TheService/Publications/2016-07-27SASLDP201617.pdf

Turner J; Coster J; Chambers D; Cantell A; Phung V-H; Knowles E: *et al* (2015) What evidence is there on the effectiveness of different models of delivering urgent care: A rapid review. *Health Services and Delivery Research 3 (43)*

Turner J; Jacques R; Crum A; Coster J; Stone T; Nicholl J (2017) *Ambulance Response programme. Evaluation of Phase 1 and Phase 2* NHS England

Addendum

(provided by the Scottish Ambulance Service and included verbatim)

Scottish Ambulance Service – New Clinical Response Model Clinical Development Overview

Over the past 40 years, there has been limited change in all UK ambulance services clinical response models. Since 1974, the Scottish Ambulance Service (SAS) and the majority of UK ambulance services, have operated a clinical response model based on expert opinion, consensus and time targets. To date, there is no substantial published evidence or consensus on best practice in ambulance pre-hospital response models within the UK or internationally. In contrast, there have been significant developments in the clinical capability and professionalisation of ambulance clinicians and a range of innovations and improvements within the Scottish Ambulance Service. In November 2016, the Scottish Ambulance Service (SAS) was supported by the Scottish Government to develop and test a new model of emergency response -a New Clinical Response Model (NCRM).

Historically, SAS had three alphabetical order, time targeted response categories, Category A (Purple and Red calls) within 8 minutes, Category B (Amber and Yellow calls) within 19 minutes and Category C (Green calls) within 60 minutes. The clinical hierarchy for these was linked to the alphabetical order denoting the most acutely ill patients to Category A. A range of resources (ambulances, cars and motorbikes) were tasked to scene at the discretion of the Ambulance Control Centre (ACC) dispatch staff to balance demand and aim to meet target response times.

Early in 2016, SAS began to investigate and develop a plan for a new approach to emergency response. Proposed changes to the response model were based on clinical data and a rationale was developed and agreed reviewed by a Clinical Advisory Group (CAG) made up of internal and external subject matter experts. External critical appraisal and approval prior to submission to the Scottish Government for commissioning was essential.

SAS has a unique advantage over other UK ambulance services as it has had Electronic Patient Records (EPR) for more than ten years. This substantial clinical data set has enabled a deep understanding of pre hospital patient interventions and outcome which provided the necessary evidence to inform and develop the NCRM.

An internal NCRM project group was established to analyse and review the clinical data and develop a proposal for change regarding improving the identification of patient’s condition and the allocation of resources. The NCRM project group reviewed 475,490 historical records from the Medical Priority Dispatch call handling triage System (MPDS). A range of data was triangulated: MPDS codes were analysed alongside internal clinical data to identify patient acuity levels, clinical interventions and outcomes. This enabled the development of a new proposed acuity level/prioritisation system associated with each MPDS code. In addition, resources were targeted in a different way: by practitioner skill level and matched resource type to meet the clinical need of patients. In essence, the aim was match and send the most highly skilled clinicians with technology and transportation to clinical priority identified.

Further understanding of the specificity of the clinical interventions and outcomes associated with MPDS codes was necessary. Therefore, in partnership with the CAG the NCRM project group developed a 10 point data analysis capture tool (Appendix A) applied to the data set. This analysis provided detail of each patient’s pre-hospital clinical acuity level by MPDS code.

SAS and CAG jointly developed and agreed the criteria for assigning acuity levels to codes for the new model. Data regarding cardiac arrest rates and interventions carried out on specific patient groups were prioritised. (Appendix A). Analysed through this matrix there was clear evidence from the data that the out-dated SAS response model was over prioritising patients as immediately life threatening for some MPDS codes and under categorising for others. The associated risks with over and under categorising of patients in the old model became the primary driver for developing a safer more efficient model to deliver improved patient outcomes.

At the time of development there were 1254 dispatch codes directly linked to the response of public 999 calls.

Emergency 999 callers are reviewed through an algorithm of questions by a non-clinical call handler using the MPDS triage system. The system subsequently generates a dispatch code by taking account of the answers to the questions during the triage process. There are 32 chief complaint categories, each with their own subset of NCRM dispatch codes and acuity levels (Appendix B).

The new model development was based on data analysis from 12 months of data, totalling 475,490 incidents of which there were 334,141 public 999 incidents and 141,349 Health Care Professional incidents.

The review focused on identifying the patient’s clinical acuity level and the planned care pathway/journey to definitive care. For example, a patient requiring resuscitation needs a multi responder team at scene, as quickly as possible, to optimise their chance of survival.

This is in contrast to a patient with an acute in-hospital pathway requirement for a timely response and diagnosis at scene to determine the care pathway and provide timely transport and access to definitive care.

Re-Categorising in this way enabled the NCRM project team to propose a new response hierarchy model. Developing the previous three tier A, B, C response to a further defined five tier coloured response hierarchy (Appendix B). This more detailed stratification of clinical acuity offered the ability to deliver a much more detailed and specific response hierarchy. The new system could now operate in a different way matching for the first time, patient clinical need, clinical skill and resource type.

Several key points were identified from the development of the NCRM;

- The systematic review of the evidence determined the rationale for change for SAS to challenge the old response category categorisation and for the first time with a robust evidence base.

- The evidence suggested that with further definition and earlier identification, a reduced number of patients who have a defined need for resuscitation/early intervention could be more accurately matched to the highest priority response.

- Earlier identification of patients in cardiac arrest or peri-arrest situation and triaged them into the purple response category. This change enabled patient groups with a high incidence of cardiac arrest to be identified from the outcome/intervention trends in the data. Historically, in the old model the patients arrest status was potentially masked by their chief complaint code for example seizure, traumatic mechanism, or unconscious.

- The NCRM specific focus on identifying high acuity patients also enabled a whole system review of the responses to patients who did not present with an immediate need for resuscitation. A key example of this change was the re-triage of previous high priority Category A calls. This is best demonstrated by the new re-triage of patients with chest pain incidents. From the systematic review, a lower rate of cardiac arrest was identified in the new model and a consistent volume of ST Elevation Myocardial Infarction (STEMI). Reconfiguring the system with the ability to send the correctly matched resource first time and consider the patients' entire pathway has the potential to further improve our efforts for safe and effective care.

The ability of the whole system to have insight and analyse patients need enables the NCRM approach to have the ability to change the culture of sending a resource to "stop the clock".

This approach underpins the policy drivers of Realistic Medicine https://beta.gov.scot/publications/practising-realistic-medicine/) 2017 and the national review of targets and performance measures to re-focus on outcomes for patients (http://www.gov.scot/Topics/Health/Quality-Improvement-Performance/Review-Targets-Indicators)2017.

The NCRM approach is dedicated to ensuring that the patients receive a clinically, timely, matched response to meet the clinical requirement, including multiple resources where necessary for potential resuscitation incidents and conveying resources.

Since the implementation of the NCRM in November 2016, a continuous review process has been place using time series data to ensure all decisions for change are data driven. SAS has standardised the way in which data and decisions rules are used to understand variation in practice. The Healthcare Data Guide (Provost and Murray 2011) Appendix C

Appendix A

Measure 1	Transported to hospital
Measure 2	Airway Intervention
Measure 3	Breathing Intervention
Measure 4	Circulation Intervention
Measure 5	Drugs Administered
Measure 6	NEWS Score +4
Measure 7	Pre-Alert
Measure 8	STEMI
Measure 9	FAST+ Stroke
Measure 10	Cardiac Arrest

Appendix B

New Clinical Response Model Hierarchy with Outline Principles
Purple Response Category ▪ Highest response priority ▪ Cardiac arrest rate over 10% ▪ Respond with closest resource ▪ Paramedic attendance essential ▪ Minimal of three responders to scene + double crewed ambulance if not in that response ▪ Consider partner agencies to support response
Red Response Category ▪ Second highest response priority ▪ Cardiac arrest rate >!% and defined need for resuscitation ▪ Response with the closest resource + double crewed ambulance if not that response ▪ Paramedic attendance essential
Amber Response Category ▪ Third response category ▪ <1% cardiac arrest rate ▪ Defined need for acute pathway care ▪ Response with the right resource – emergency transporting ambulance ▪ Paramedic attendance preferred
Yellow Response Category ▪ Fourth response category ▪ <1% cardiac arrest and no defined acute pathway care ▪ Response with the right resource – ambulance for defined hospital need and PRU for potential alternative pathway care
Green Response Category ▪ Fifth response category ▪ Exclusion of above categories ▪ Potential for additional clinician led telephone triage ▪ Face to face assessment when required

Appendix C

There are four key concepts for determining non-random 'signals':

Rule 1: Shift Six or more data points either all above or all below the median

Rule 2: Trend Five or more consecutive points in the same direction (upward/downwards)

Rule 3: Number of Runs Too few or too many runs, or crossings, of the median line

Rule 4: Astronomical Data Point Undeniably large or small data point

http://www.med.unc.edu/cce/files/education-training/The%20run%20chart%20a%20simple%20analytical%20tool.pdf

Ref 1 NCRM Phases paper
Ref 2 NCRM Call Volume and Clock Start paper

New Clinical Response Model Phases

Phase 1 of the new clinical response model (NCRM) went live on 23rd November 2016. Phase 1 consisted of changes to acuity levels associated with codes and the configuration of pre-determined resources on the command and control system to guide dispatchers on the type of resources (s) required to meet the clinical needs of the patient.[1] This was supported by a change to any standard operating procedure that impacted dispatch decision making to ensure that staff had clear guidance on dealing with all eventualities, for example, what to do if an ILT call when a crew were already on another call, on a rest period, at the end of their shift etc.

Phase 1 Developments Timeline

- NCRM Code Acuity Level Changes 23rd November 2016
- Pre Determined Attendance assigned to codes in C3 System 23rd November 2016
- Updates to Standard Operating Procedures 23rd November 2016

Phase 2 of the NCRM consisted of enhancements in technology to aid and support dispatch decision making, and a number of process improvements aimed at reducing waste, ensuring the right clinical resources to patients first time and improving the early identification of immediately life threatening patients and patients in cardiac arrest to ensure early initiation of bystander CPR and intervention using public access defibrillation to increase survival rates.

Phase 2 Developments Timeline

- Enhanced Defib Module /PAD Mapping 29th March 2017
- First Responder Module 15th May 2017
- ProQA MPDS Triage System Upgrade Version 13 14th June 2017
- C3 Nexus ELAN 14th June 2017
- Pre-Entry Questions Module (PEQ 1) 28th August 2017
- PEQ 2 and Dispatch on Disposition (DoD) 10th October 2017

1. Enhanced Defibrillator Module

The enhanced defibrillator module is an enhancement within the C3 command and control dispatch system that allows public access defibrillators to be entered and mapped onto the system and configured to be highlighted only if the dispatch code is appropriate and the public access defibrillator is at a suitable distance from incident. This development enables the identification of public access defibrillator against location and proximity of patients in cardiac arrest and where appropriate highlights that there is a public access defibrillator within proximity, allowing the call handler to signpost bystanders to the location defibrillator.

2. First Responder Module

The First Responder module is an enhancement within the C3 Command and Control dispatch system that aids and supports better dispatch decision making. This module allowed the system to be configured with set criteria for dispatching first responders, including dispatch codes and age criteria suitable for first responders as well as distance of first responder from location of incident. For any incident the system assesses the call against suitability of code, age and proximity and highlights any suitable incidents to the dispatcher for first responder allocation. Prior to the implementation of the first responder module, the dispatcher had to manually assess the call suitability for first responders which was time consuming and inconsistent.

3. C3 Nexus ELAN

The C3 Nexus ELAN upgrade to the system was a development that enabled auto population of address locations of incidents on the system taking account of the location of the callers, to save vital time in address location verification.

4. Dispatch of Disposition (DoD)

The Scottish Ambulance wasn't the first Ambulance Service in the UK to adopt Dispatch on Disposition (DoD). In January 2015 approval was reached from the Department of Health in NHS England (NHSE) to pilot DoD, this was the first phase of the new Ambulance Response Programme (ARP) approved for implementation in NHSE. In early February 2015, the first phase of the ARP went live in the London Ambulance Service (LAS) and the South Western Ambulance Service NHS Foundation Trust (SWAST) with the North East, South Central, West Midlands and Yorkshire Ambulance Trusts following suit in late October 2015.

The essence of DoD is that, for patients who are not presenting with an immediately life threatening condition, taking a short amount of additional time to clearly understand the clinical needs of the patient will result in a better match of resource to need.

SAS went live with DoD on 10th October 2018. Prior to the implementation of DoD, practice was broadly centred around allocating the nearest resource on confirmation of address, in pursuit of delivering against an 8 minute response time target. The old dispatch on address model meant that dispatchers were blindly sending resources to incidents where the patient's condition and acuity level had not yet been established, often resulting in patients with non life threatening conditions receiving a quicker response than patients with immediately life threatening symptoms in need of urgent clinical intervention. There was recognition that moving away from dispatching on address and taking the time to understand the condition of the patient before sending a resource would mean a potential delay in allocation of resources to all patients and as such a need to have a system in place that would allow call handlers to identify cardiac arrest and immediately life threatening patients at the earliest possible stage in the call cycle.

Purple ILT and Cardiac Arrest response times has remained statistically stable following the implementation of dispatch on disposition due to the ability to identify a high proportion of these patients at a very early stage in the call (PEQ stages). PEQ is successful in identifying around 35% of all ILT patients at PEQ stages, patients in the red ILT category normally take longer to establish the presence of ILT symptoms, requiring further triaging before confirmation. Prior to DOD, dispatchers could blindly dispatch on address before knowing the acuity of the patient which meant that resources were often allocated and on route to a patient before the performance clock started. Since DOD, the dispatcher allocates a resource on confirmation of the condition of the patient unless final coding hasn't been reached at 180 seconds, in which case the call presents to the dispatcher for allocation of a resource. The performance clock start has not been changed in line with dispatch on disposition to reflect the new process of dispatching a resource once the patient's condition is known. This measure of response time is currently under review with Scottish Government.

5. Pre Entry Question Module (PEQ)

The PEQ module within the command and control system was introduced at the same time as DoD to ensure early identification of patients in Cardiac Arrest and patients with immediately life threatening symptoms to allow rapid dispatch of resources to those most in need. Pre Entry Question 1 "Is the patient breathing" was implemented in August 2017, followed by Pre entry Question 2 "Is the patient awake" on 10th October alongside DoD.

The ability to identify patients in cardiac arrest at an earlier stage in the call cycle is a key enabler for delivering the out of hospital cardiac arrest strategic aims of saving more lives by initiating quicker hands on chest/CPR telephone advice instructions and signposting bystanders to public access defibrillator where available.

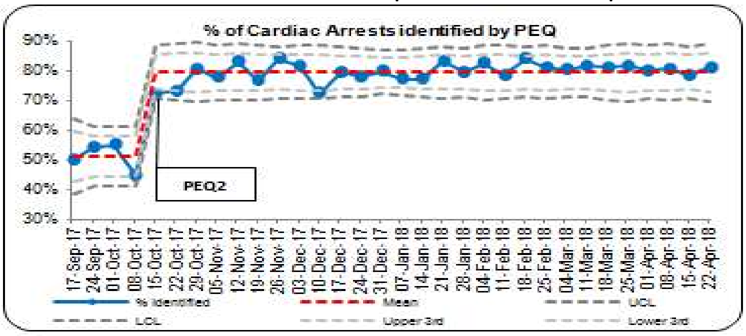
Phase 3 NCRM

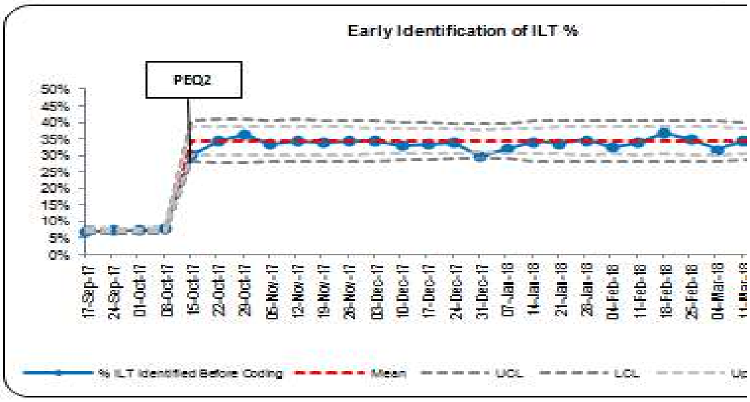
There is ongoing internal audit, monitoring and review of NRCM developments and continuous refinement to the system.[1] Phase 3 of NCRM is primarily about refining the model to further improve patient outcome and care. All changes are evidence led, clinically assessed and modelled for forecast impact prior to implementation.

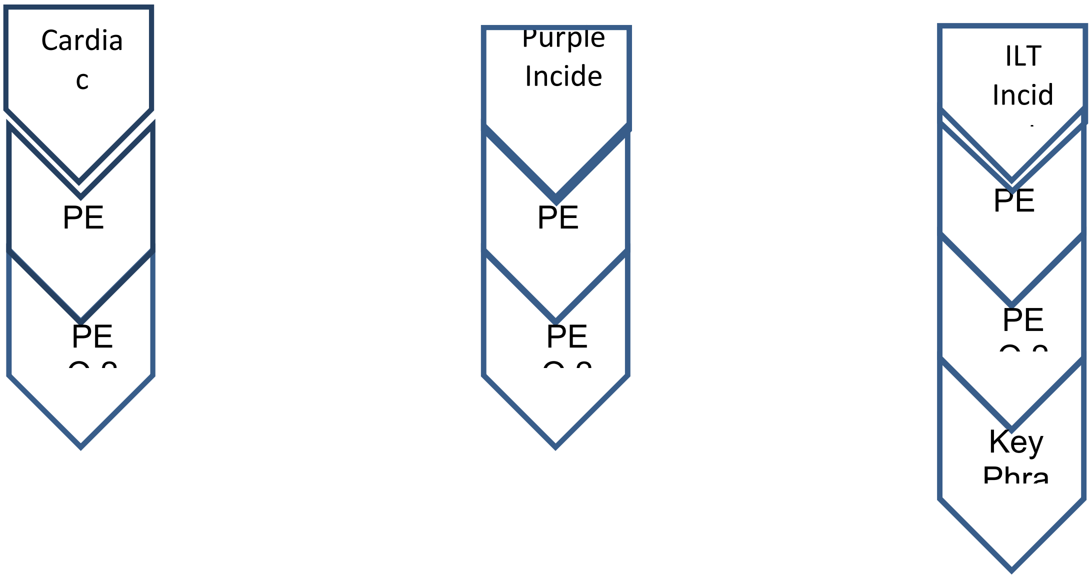
In April 2018 'Key Phrases' were introduced as part of phase 3 of the project to further increase the early identification of ILT patients. The Key phrase module in the command and control system works by identifying another group of ILT incidents from the specified words entered in the 'What's the Problem' field. If a potential ILT is identified through key phrases it will be sent to dispatch for immediate action.

Early Identification of Immediately Life Threatening Patients

The data used to monitor the impact of the implementation of Pre Entry Question 1 and 2 and Key phrases indicates that there has been a significant improvement in the early identification of cardiac arrest patients and ILT patients.







Phase 3 Developments Timeline

1. Key Phrases	9th April 2018
2. Yellow Basket Split	Work in progress
3. HCP Card 35 Triage	Work in progress

Ref 1 – NCRM Background Context paper

<u>Call Volume and Clock Start</u>

There have been a number of key areas outlined in this document that have an impact on the ability to compare colour categories survival rates and response times in January 2016 to post NCRM[1] January 2017 and 2018.

1. **Response Time Comparison**

As well as the changes in process for DOD[2] there are 2 other key specific areas within the data that makes response time in January 2016 non comparable with post NCRM data in January 2017 and January 2018.

- Volume of Clinical Support Desk Upgrades
- Volume and associated performance of HOSRED calls (Immediately life threatening hospital calls)

NCRM was live in January 2017 and 2018 however the adverse weather and presence of flu in the system during the winter of 2017/18 has had a significant impact on the number of immediately life threatening patients in the cardiac arrest and breathing/respiratory categories.

<u>Clinical Support Desk (CSD) Upgrades</u>

CSD upgrades are calls that have undergone secondary triage by a clinical adviser and subsequently upgraded to a higher acuity level.

There was an increase in the number of incidents in January 2017 and 2018 which have been upgraded to ILT (purple and red) by clinical advisors. These incidents will have been initially coded as a lower acuity (amber, yellow, green) but were upgraded as a result of a review by a clinical advisor. The call started time for these incidents is maintained from the original call started time in the lower acuity call, meaning that these incidents will have a longer time from call started to first allocated time and a longer time from call started to first arrived scene time. The increasing number of these incidents means they are having a larger effect on the average allocation and response times for red calls in 2018 than previous years.

Table 1 - Number of Incidents (CSD and Non CSD)

	Jan-16		Jan-17		Jan-18	
	Non CSD	CSD	Non CSD	CSD	Non CSD	CSD
Purple	414		703	1	910	6
Red	12,819	28	3,956	211	5,089	371
Amber	1,554		8,930		11,623	
Yellow	27,154	1,030	27,271	1,608	27,004	1,527
Green	3,022	769	2,970	1,027	3,515	458
Unknown	1,754		1,911		2,368	
Sum:	46,717	1,827	45,741	2,847	50,509	2,362

Table 2 - Time Call Started to First Allocated (minutes) for CSD and Non CSD Incidents

	Jan-16		Jan-17		Jan-18	
	Non CSD	CSD	Non CSD	CSD	Non CSD	CSD
Purple	1.48		1.64	14.7	1.95	36.74
Red	2	7.29	2	15.52	3	37.16
Amber	2.46		2.77		11.59	
Yellow	3.30	8.56	4.27	13.55	13.90	47.03
Green	3.19	13.33	3.00	15.86	35.37	46.67
Unknown	1.72		1.71		3.36	
Sum:	2.75	10.48	3.62	14.49	12.01	45.28

Table 3 - Time Call Started to First Arrived Scene (minutes) for CSD and Non CSD Incidents

	Jan-16		Jan-17		Jan-18	
	Non CSD	CSD	Non CSD	CSD	Non CSD	CSD
Purple	8.58		8.53	21.57	8.64	60.31
Red	9	25.67	9	30.69	11	52.51
Amber	14.17		14.13		24.14	
Yellow	14.75	34.30	16.47	40.43	28.07	68.52
Green	34.72	58.51	37.17	65.00	81.11	73.05
Unknown	7.83		8.26		5.68	
Sum:	13.14	43.30	15.32	46.98	24.89	66.31

Emergency Hospital Transfers (HOSRED)

Prior to the introduction of NCRM the majority of emergency hospital transfers were coded as red calls. If the call originated from a hospital which held a defibrillator and appropriately trained staff then a hospital 'resource' could be allocated to the job. The effect of this was that these resources were allocated, arrived on scene and cleared within a few minutes (Average response time of 3.76 minutes from the call started time). An ambulance resource would also be deployed to these incidents but would not be first on scene. Since NCRM implementation, HOSRED has ceased to exist meaning that hospital resources no longer count towards response time performance for those ILT calls that originate in a hospital. In January 2016 there were 984 HOSRED calls with an average response of 3.76 minutes, compared to none in 2017 and 2018 and as such 2016 response times are artificially lower than January 2017 and 2018 where only ambulance response time counted towards performance.

Table 4 - Number of Incidents (HOSRED and Non HOSRED)

	Jan-16		Jan-17		Jan-18	
	Non-HosRed	HosRed	Non-HosRed	HosRed	Non-HosRed	HosRed
Purple	414		704		916	
Red	11,863	984	4,167		5,460	
Amber	1,554		8,930		11,623	
Yellow	28,184		28,879		28,531	
Green	3,791		3,997		3,973	
Unknown	1,754		1,911		2,368	
Sum:	47,560	984	48,588		52,871	

Table 5 - Time Call Started to First Allocated (minutes) HOSRED and Non HOSRED incidents

	Jan-16		Jan-17		Jan-18	
	Non-HosRed	HosRed	Non-HosRed	HosRed	Non-HosRed	HosRed
Purple	1.48		1.66		2.17	
Red	1.68	1.18	2.58		5.32	
Amber	2.46		2.77		11.59	
Yellow	3.50		4.79		15.71	
Green	5.69		7.07		39.83	
Unknown	1.72		1.71		3.36	
Sum:	3.10	1.18	4.27		13.60	

Table 6 -Time Call Started to First Arrived Scene (minutes) HORSED and Non HOSRED incidents

	Jan-16		Jan-17		Jan-18	
	Non-HosRed	HosRed	Non-HosRed	HosRed	Non-HosRed	HosRed
Purple	8.58		8.55		8.98	
Red	9.57	3.76	10.39		13.70	
Amber	14.17		14.13		24.14	
Yellow	15.47		17.82		30.29	
Green	48.87		55.82		76.89	
Unknown	7.83		8.26		5.68	
Sum:	14.55	3.76	17.21		26.85	

Comparing Survival Rates by Colour Category

The volumes of ILT incidents in January 2017 and 2018 are significantly lower than 2016 and the acuity level of the patients in the ILT categories in 2016 is not comparable to 2017 and 2018. In 2016 there were a significant number of patients in the ILT categories that were at very low risk of cardiac arrest/not ILT. The NCRM was designed to ensure that only those patients who were in cardiac arrest or high risk of cardiac arrest remained in the immediately life threatening categories and those at lower risk were moved into appropriate categories in Amber, Yellow or Green.[1] This also means that patients in the Red and Purple categories in 2017 and 2018 are much sicker than in 2016. For example, see Table 7 – In the red breathing difficulties category there are 2587 incidents in 2016 compared to 540 in 2017 and 947 in 2018.

As part of phase 1 of the New Clinical Response Model the number of MPDS codes which make up the 'breathing difficulties – red' group reduced from 7 in January 2016 to 2 in January 2017 and 2018. This was based on the evidence collected and analysed by the Clinical Team. 4 of the codes which were moved out of the 'red' basket were moved to Amber and 1 was moved to purple. Those that were moved to Amber were moved because there was clinical evidence that those patients had a very low risk of cardiac arrest and those that were left in the red category or moved to purple had a high or very high risk of cardiac arrest.

Table 7 – Incident Summary

Variable	Cross tab	Jan-16	Jan-17	Jan-18
		Count	Count	Count
CAD Incident Summary	Cardiac Arrest (Purple)	397	556	723
	Breathing Difficulty (Red)	2,587	540	974
	Stroke (Amber)	925	1,217	1,279
	Falls (Yellow)	4,023	4,324	4,746
	Not Specified	40,612	41,951	45,149
	Sum:	48,544	48,588	52,871

Those codes where there was evidence of a low cardiac arrest rate that moved from Red in 2016 to Amber in 2017 include:

- 06D02 – difficulty speaking between breaths – 1,480 incidents in 2017
- 06D02A – difficulty speaking between breaths – known asthma – 611 incidents in 2017

January 2018 Pressures

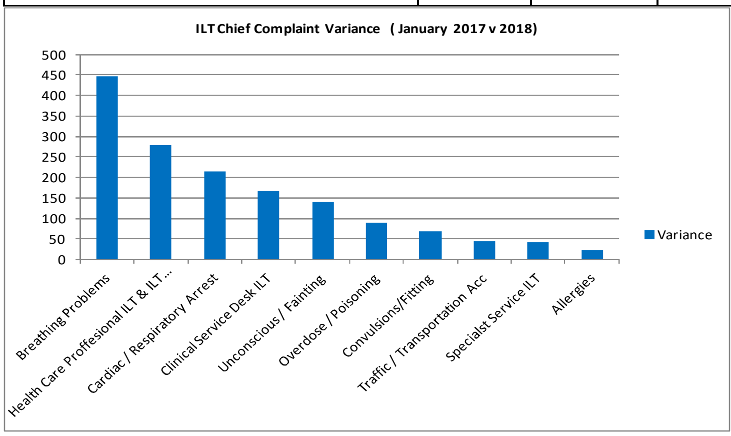
There were a number of internal and external pressures within the system since mid November, throughout December and in January, which also had an impact on response times for patients, including increased demand, flu, all time high hospital turnaround times and adverse weather conditions. The impact and pressures of the flu and cold weather were witnessed across the wider healthcare system.

Emergent Demand for SAS was 4.55% higher this January 2018 in comparison to January 2017 and Immediately Life Threatening Demand was up 32.35%.

	Jan-17	Jan-18	Variance	Variance %
Incidents Responded to	54594	56077	1483	2.72%
Emegency Incidents Responded to	42478	44411	1933	4.55%
ILT Incidents Responded to	4729	6259	1530	32.35%
ILT as % of Emergency	11.1%	14.1%		2.96%

The largest increase in ILT was from patients presenting with breathing problems, Healthcare professional ILT calls and Cardiac Respiratory Arrest calls. The table below contains the top 10 chief complaints with the highest increase.

Chief Complaint Group Desc	January 17	January 18	Variance	Variance %
Breathing Problems	540	988	448	83.0%
Health Care Proffesional ILT & ILT Transfer	194	474	280	144.3%
Cardiac / Respiratory Arrest	548	763	215	39.2%
Clinical Service Desk ILT	210	376	166	79.0%
Unconscious / Fainting	683	823	140	20.5%
Overdose / Poisoning	172	261	89	51.7%
Convulsions/Fitting	901	968	67	7.4%
Traffic / Transportation Acc	329	374	45	13.7%
Specialst Service ILT	0	41	41	
Allergies	144	167	23	16.0%



Ref 1 – NCRM Background and Context paper

Ref 2 – NCRM Phases