Fire and Rescue operational effectiveness: the effect of alternative crewing patterns

Abstract

This paper assesses the operational effect of changes to crewing patterns in a UK Fire and Rescue Service (FRS). It is set in the context of two main drivers of change in the UK fire service; decreasing demand in the UKFRS and reduced budgets. A novel framework for assessing operational effectiveness was developed within a longitudinal mixed method design that was used to integrate two years of existing operational data with data from interviews with firefighters.

Results show changes in crewing patterns can be implemented without a substantial quantitative impact on overall operational effectiveness but there is the potential for a negative effect wider service capacity. In addition the development of assessment for operational effectiveness in the FRS provides the ability to inform decision making in an authentic way that allows stakeholders to have confidence in the outcomes, whilst being timely and not too complex or costly to evaluate. This approach provides an important contribution to practice in terms of the assessment of public services in a time of challenging reform and demonstrates that alternative crewing patterns to better match demand can be implemented provided careful consideration is paid to wider considerations.

Keywords

Overall Equipment Effectiveness, public sector, performance measures, emergency logistics

1 Introduction

A ‘performance revolution’ (Neely, 1999) has transformed performance measurement for public services over the last thirty years (OECD, 2004). Against a background of public sector reform (Public Accounts Committee, 2011), UK Fire and Rescue Services have considered new ways of working, including changes to crewing patterns to fulfil their remit of prevent, protect, and respond (Knight, 2013).

In this paper we examine the measurement of operational activities, in the context of crewing change within four fire stations of the Counties FRS (The Counties is pseudonym to protect anonymity) . The change was from a 2-day, 2-night, 4-off (2:2:4) rota to an alternative crewing arrangement (ACA) involving longer spells of attendance by fewer firefighters. We offer a model of operational effectiveness (OE) that aims to meet the OECD (2004) criteria of being reliable for stakeholders, timely and not too complex or costly to evaluate.

1.1 UK FRS and the Counties - demand and budgets

Demand from all incidents experienced by the UK FRS has reduced by 46% since 2002 (Department of Communities and Local Government, 2013[a])[1]. Thus the Counties has explored more efficient ways of managing capacity including ACAs which allow fire stations to be staffed by whole-time firefighters but at a lower level of capacity than conventional (2:2:4) fire stations. ACAs offer a potential alternative to reducing capacity without a reduction in responsiveness along with a cost saving of up to £370,000 per year per station (The Counties, 2012[c]). In 2002 there were 31,761 wholetime (FTE) firefighters in the UK but this reduced by approximately 27% (22,957) by 2017 (Home Office, 2018). Over the same period the number of retained firefighters (FTE) grew between 2002 and 2011 but by 2017 numbers had fallen to below that of 2002 (10,092) (Home Office, 2018).

[1] Official statistics (Home Office, 2017[a]) report a 37% increase in fire deaths in 2017 compared to the previous year, however these figures include the 71 fatalities from the Grenfell Tower fire. Anecdotally, the UK national press and the Fire Brigades Union (The Guardian, 2017; Fire Brigades Union, 2017) have reported an increase in fire related fatalities, however confirmation as to whether this is a continuing trend is not possible within the time frame of this paper.

The Counties ACA involved firefighters living in stations for up to five continuous days and nights, with 12-hours 'down time' in each 24-hours. The ACA required 50% fewer personnel than traditional 2:2:4 shift patterns but with no reduction in geographic cover (The Counties, 2012[a]) and is an attractive crewing arrangement for fire stations that have lower call-out demand. A 2:2:4 system requires seven personnel to crew a watch on a one pump (fire engine) wholetime station, including the Junior Officers (JOs), with four watches per station. Retained fire stations are crewed by on-call firefighters who live or work in the local area. These represent a skilled flexible workforce to fill gaps in capacity but may not be sufficient for medium demand stations.

1.2 Performance measurement in the UK Fire and Rescue Service

The UK FRS annual key performance indicators, governed by the Department of Communities and Local Government (DCLG), exemplify the new public management (NPM) ethos (Carvalho *et al.*, 2006). NPM focuses on efficient use of resources, performance monitoring and greater internal and external accountability to address the requirements of a wide range of stakeholders (Kloot, 2009).

Murphy and Greenhalgh (2013) suggest that the maturity of FRS performance management regime falls between the 'data rich' and 'intelligent data'. Whilst the UK FRS continues on a path towards self-regulation, Theresa May (2016), the then Home Secretary, highlighted the continuing difficulties in governance and scrutiny of the FRS due to the lack of clear auditing practices and the limited available data between services and over time. In parallel to the 2012 reforms carried out within the Police Force, May pushed for transparent, publicly available data to allow for comparisons across FRS (May, 2016).

A UK FRS toolkit facilitates self-reflection, peer auditing, and strategy development (Chief Fire Officers Association, 2012) and 'a robust process for assessing operational performance'

(Local Government Association, 2015 pp.4). The Operational Assessment (OpA) and Fire Peer Challenge aims to highlight key areas of concern, enable effective service delivery, encourage both continuous improvement and enhance accountability and a ‘whole system’ approach (Chief Fire Officers Association, 2012; Local Government Association, 2015). However, the toolkit relies heavily on subjective qualitative benchmarking thus preventing rigorous empirical comparisons between services and also runs the risk of ‘gamification’ (Downe *et al.*, 2017). Chief Fire Officers are also able to select their reviewers creating the potential for bias and potentially compromising the credibility of recommendations. There are also questions about how ill-defined concepts, such as operational/organisational effectiveness, are measured and terms such as ‘uses available data to benchmark performance and support improvement’ (Local Government Association, 2015 pp. 13) are open to interpretation.

Murphy and Greenhalgh (2013) call for three principles that that could be employed to map more rigorously and accurately FRS performance. Firstly, national and local indicators would allow for empirical comparisons over time and across services. These should be ‘quality assured, robust and accredited standards and benchmarks publicly available from an independent host’ (Murphy and Greenhalgh, 2013 pp.231). Secondly, the analysis of key documents such as the Integrated Risk Management Plans to identify areas of importance to each FRS. Thirdly, the commissioning of academic and operational research will help move the FRS towards a more mature performance monitoring regime and stimulate dissemination of good practice; however, this can only be achieved once there are clear measures to allow for meaningful comparisons to be made.

Despite these recommendations, there has been no change in the way performance is measured and reported within the FRS. To address absence of an established framework to assess operational performance within the FRS, this research offers a method to assess the operational effect of ACA.

1.3 Operational Effectiveness in the FRS

Little rigorous research has been conducted on operational effectiveness within the UK FRS, with only a small number of studies focusing on response times (Taylor, 2016; Bateman *et al.*, 2016; Jaldell, 2005). In terms of ACAs assessment of operational effectiveness is needed because of the potential to impact on operations of the stations by condensing the time available for training, preventative work, and routine maintenance. In the case of this ACA, extended hours at work may also increase fatigue, impacting on the way firefighters can carry out their duties.

Reduction in demand for emergency incidents offers a rationale for lowering capacity. However, public expectation is that the FRS will maintain and improve responsiveness to emergencies. Response times need to be maintained and geographic coverage must be upheld in line with Fire and Rescue Services Act (2004). Fewer emergency incidents alongside demands to maintain coverage and responsiveness presents a challenge to FRSs when considering reducing staff.

Operational effectiveness (OE) is defined as those activities that enable an organisation to a) better utilise its resources, b) better implement its processes and c) achieve its mission and objectives (Porter, 1996). OE is a central feature of the continuous improvement of functional performance, commensurate with the FRS OpA and Peer Challenge.

In manufacturing, capacity it is usually linked to output. For services, capacity is more difficult to define and is a balance between human and physical resources. In service industries fluctuating demand requires adjustments to capacity. Seasonality and variability of demand is a particular challenge for the emergency services where the ability to respond is paramount (Richie and Walley, 2016).

Previous work has attempted to quantify FRS performance in terms of input and output to fill the gap in efficiency and productivity measurement within this sector (Jaldell, 2005). Response times were used to assess the impact of station closures within London Fire Brigade, following Central Government's cuts to public services (Taylor, 2016), but poor variability in response times limits the use of this variable as a measure of effectiveness (Carvalho *et al.*, 2006; Jaldell, 2005). However, the strategic importance of response times for key stakeholders indicates a need to include these times as part of any measure of operational effectiveness.

The Fire Services Act (2004) places greater emphasis on the documentation of clear national and local priorities and objectives for FRS authorities. The government direction for FRS in England is set in the 'Fire and Rescue National Framework for England' *(*2012) stating it; '..sets out high level expectations. It does not prescribe operational matters. These are best determined locally by fire and rescue authorities.'.

The priorities are:

'1. identify and assess the full range of foreseeable fire and rescue related risks their areas face, make provision for prevention and protection activities and respond to incidents appropriately

2. work in partnership with their communities and a wide range of partners locally and nationally to deliver their service

3. be accountable to communities for the service they provide'

(Department for Communities and Local Government, 2012[b] pp.7)

Operational performance measures should address Priority 1 in terms of prevention, protection and responding, and Priority 3 by ensuring a clear accountability providing a demonstration of Priorities 1 and 2.

Conventionally, performance objectives focus on quality, cost, flexibility, speed, and dependability (Slack *et al.*, 2010). Performance measures can then be designed to ascertain how closely an organisation is meeting those objectives. As part of the evaluation of the Counties operational effectiveness, a list of current performance measures was compiled through discussion with data analysts from the service and follows similar measures of a comparable FRS (Bateman *et al.*, 2016). This approach was based on the development of Overall Equipment Effectiveness (OEE) (Nakajima, 1984) which combines the operation, maintenance, and management of equipment to measure effective use of resources. Endorsement from the UK government (DTI, 2000) has led to wide acceptance of OEE within UK manufacturing. It measures the effectiveness of equipment via the three indicators of performance (*p*), availability (*a*), and quality (*q*) and is used to track and trace improvements or decrements over time.

OEE has been applied in the road freight industry (Simons *et al.*, 2004) and using OEE in the FRS involves viewing a fire appliance with its crew as a piece of operational equipment. Within the service industry the areas of availability, performance, and quality are highly relevant to service delivery and work with a comparable FRS sufficiently positioned FRS performance measures within the OEE framework (see Table 1) (Bateman *et al.*, 2016).

Table 1. Summary of suggested operational effectiveness measure (from Bateman et al., 2016)

Conventional OEE	Possible measures	Comments on rigour and usability of data
Performance	Turnout times Attendance times	Probably do not vary much Too incident specific and dependent on geography?

Availability	Sickness (neg) Stand down times (neg)	Is availability too narrow a measure for the fire service?
Quality	Commendations Training competency Breathing Apparatus tests Vehicle accidents (neg) Critical equipment fails (neg) Discipline cases (neg) Personal injuries (neg) Equipment maintenance	All relatively indirect and either look at preparedness or HR issues Maintenance is mainly performed by the firefighters and interviewees indicated that this measure seems to indicate a well-run station

Overall operational effectiveness involves complex interactions between different processes and the isolation of data can miss these interactions when using OEE (Muchiri and Pintelon, 2008). Priority 1 for the FRS is to '*identify and assess the full range of foreseeable fire and rescue related risks their areas face, make provision for prevention and protection activities and respond to incidents appropriately.'* This suggests an interaction between different areas of the service and more than simply the fire appliance with its crew. When considering this alongside the suggestions from FRS stakeholders, we proposed an alternative model assessing *Responsiveness, Preparedness,* and *Personnel Measures* (see Table 2).

Responsiveness (turnout and attendance times) has diminished in significance as a performance indicator being removed from the Audit Scotland publications in 2005 (Carvalho *et al*., 2006). Target times do not reflect the individual nature of each incident, such as road conditions, location of the incident, time of day etc. Whilst response time is still used as an indicator in England and Wales, lack of variation can be seen with the Counties achieving 94-98% on or under target. Whilst targets linked to these data are a key part of the Counties strategic plan, they should not be relied upon as the only measure of ACA.

Preparedness and personnel measures may give a greater degree of variability over time and across stations tracking any subtle impact of the ACA.

Measurement of performance is at the station level, both before and after implementation of the new shift system allowing for the nuances of ACA introduction at individual stations to be explored with the context of each in mind. Station level performance data can then be compared to the service as a whole, providing an interpretative framework especially for new fire stations where pre-ACA data is unavailable.

The Overall Effectiveness of Fire Operations framework used in this study is outlined in Table 2 and was adapted from Bateman *et al.* (2016) to provide a more appropriate framework for the evaluation of ACA. Quality was translated to preparedness with the focus towards the organisation being prepared to meet the demands of its core business activity (attending and dealing with emergency incidents). It was broadened to include more than just personnel measures, e.g. equipment failures.

Table 2. Proposed Overall Effectiveness for Fire Operations (OEFO) framework (adapted from Bateman et al., 2016)

Conventional Overall Equipment Effectiveness (OEE)	**Overall Effectiveness of Fire Operations (OEFO)**	**Related Performance Measures** (neg) denotes a reduction in incidents relates to improvement in performance
Performance	Responsiveness	Turnout times Attendance times
Availability	Availability	Sickness (neg) Pump availability (neg)
Quality	Preparedness	Equipment maintenance Training competency Critical equipment fails (neg) Breathing Apparatus tests Commendations Vehicle accidents (neg) Discipline cases (neg) Personal injuries (neg)

2 Method

In addition to the collection of KPI data outlined the framework (Table 2), 45 individuals were interviewed (8 Watch Managers, 7 Crew Managers, 30 Firefighters), from a total population of 64 at four ACA fire stations. Semi-structured interviews were conducted as part of a wider study at three time points; pre-changeover, at transition and 12-months post change. The focus of this paper is on the quantitative measure of operational effectiveness but relevant commentary from the interviews on performance have been included. The data collected may be affected by factors unrelated to ACA, as well as influenced by ACA. Interviews were used to discuss participants' thoughts around operational performance following the change to ACA. Questions within the interview schedule related to thoughts around the impact on performance at incidents and day to day duties. These interviews took place prior to the collection and analysis of the performance data and performance was explored in a general sense.

2.1 Pilot study

Identification and exploration of the KPI data as part of a pilot study assessed the data's suitability for the evaluation of ACA and helped refine the OEFO framework for the full study. The pilot comprises data for all identified KPI's within the OEFO framework for a three-month period in 2012 (pre-ACA) and 2013 (post-ACA) for Station D. After discussion with data analysts at the Counties FRS the months May-July were selected for greatest yield of data as this was felt to be the busiest period in terms of operational demand.

Pilot Study Findings

Responsiveness. A three-month data collection period was not long enough to gather valid findings for the responsiveness domain as there were only 10 life critical incidents in the three-month period in 2012 and 9 in 2013. A longer time period (12-months) for both pre- and post- ACA is to be examined in the main study.

Availability. The availability of the wholetime appliance was very high at 99.5% or above. The strategic importance of wholetime appliances, may require extra resources obtained to keep these appliances available. Factors influencing the availability of appliances and any differences for retained versus wholetime could identify any impact from ACA in the full study.

Preparedness. Preparedness was assessed looking at a range of measures such as firefighter training, equipment testing and community events. Bateman et al., (2016) identified these measures as accessing the ability for firefighters to respond as expected at emergency incidents, alongside value added activities within the community. The pilot study data is outlined in Table 3.

Table 3. Pilot study measures of preparedness (Station D)

Assessment Measure	Pre-DCP	Post-DCP
Equipment testing (N° of tests) • Weekly • Monthly • Quarterly	 433 16 68	 506 65 60
Hydrant checks (N° of checks)	110	78
Community events (time in minutes)	6840	9680

Feedback from data analysts at the Counties FRS on the utility of this data indicated direct before and after comparisons over a three-month period would be crude. Hydrant checks needed to be assessed over a longer period because of a 24-month testing cycle. The number of visits to properties for Home Fire Safety Checks (HFSC's) needs to be considered alongside successful completion of such visits as successful completion checks is not always possible. Referrals from other agencies for HFSCs can also influence how many are attempted and completed. Procedural changes unrelated to ACA, such as the introduction of Redkite (safety management system for testing of operational equipment) and the use of equipment in operational situations may impact on equipment testing statistics. Qualitative data collected during interviews were used to identify mechanisms that might link changes associated with ACA or other factors on the measures of preparedness. Other factors identified in the preparedness section were either not available, infrequent events, or partially represented in another measure, for example personal injuries would show in sickness absence in availability.

Conclusion from Pilot study

The three domains of the OEFO framework (responsiveness, availability and preparedness) provide a useful measure of performance following the introduction of ACA. The final framework for assessing the operational effectiveness is summarised in Table 4, and includes the performance indicators identified from the pilot study as having the greatest utility.

Table 4. Performance measures for full study mapped to OEFO framework

Overall Effectiveness of Fire Operations (OEFO) category	**Key performance measures (unit)**
Responsiveness	Turnout times (sec) Attendance times (sec)
Availability	Sickness (days/person) Pump availability (%)
Preparedness	Training drills (nº/person) HFSC's (nº/station) Equipment tests (nº/station) Hydrant Checks (nº/station) Community events (mins/station) Health and Safety events (nº/station) Commendations (nº/person)

Following assessment of the available data, three months did not provide a great enough yield of data to be able to infer impact of ACA on performance. Therefore, the full study draws on 12-months data pre and post ACA for each measure.

3 Method for full study

3.1.Data collection method

12-months existing performance data for each of the measures in Table 4 were collected for each station pre- and post-ACA starting from each station's respective date of ACA implementation. With Station A being a new station and therefore having no pre-ACA data, data for all the Counties FRS wholetime stations in combination were collected to provide an overall mean for the service as a comparison.

Two members of the Counties FRS staff were responsible for collating the data from the relevant service databases and collating it within an excel spreadsheet for the principle investigator (PI). Each measure's data were presented as means per month, per station.

3.2. *Analysis Method*

Descriptive statistics for each of the measures were scrutinised for percentage changes between the two time-points. For explanatory purposes a difference of 0-3% was identified as minimal/no change, 4-10% was identified as a small change (positive or negative) and above 10% was identified as a large change (positive or negative).

Findings have been categorized as per the OEFO framework (Table 4); Responsiveness, Availability and Preparedness. Data were screened to ensure assumptions for parametric testing were met; data were normally distributed. Inferential statistics using *t*-tests assess whether these differences are statistically significant. Repeated measures *t*-tests compared pre and post data for Station B, Station C, and Station D. Independent samples *t*-tests compared Station A data against the Counties means. Given the potential for programmatic changes based on the findings, a Bonferonni correction for multiple tests adjusted the accepted statistical significance to $p \leq .002$.

It is important to note that a change may not be statistically significant but still be of importance to the Counties FRS. Small changes that fail to reach statistical significance have not been dismissed if it falls below the acceptable standard for the Counties FRS. Findings are discussed in line with targets laid down in the Counties Integrated Risk Management Plan (IRMP) and other strategic documents to ensure relevance to stakeholders and within the boundaries of acceptable change.

4 Findings

Table 5 provides a summary of the comparisons between pre and post ACA for the OEFO measures. Responsiveness has not seen either a strong improvement or decline across Station B, C or D. The decline for Station A can be accounted for by the indirect comparison with The Counties mean.

Availability has seen a decline for retained appliances for Stations B, C and D and the Counties as a whole. There is no retained appliance at Station A. Wholetime appliance availability has remained high and fluctuations have been small. Sickness absence has improved for Stations A, C and D. Station B has seen increased sickness absence, however, further scrutiny of the raw data has uncovered this was unrelated to ACA.

Preparedness has seen an increase in the number of drills per person, except for Station C. There has been strong decline in HFSC's for Station D, and for Station A when compared to the Counties mean. None of the stations show an overall decline or improvement in performance.

4.1 Summary of inferential findings

A total of 26 comparisons were conducted.

- 16 comparisons were non-significant ($p > .05$)
- Seven comparisons were significant at $p < .05$ but were non-significant once the Bonferroni correction was applied. Two of these comparisons were approaching the corrected *p* value of .002.
- Three comparisons were significant at $p < .002$ and all related to retained appliance availability: Retained availability for the Counties as a whole does not appear in the summary table.

Non-significant comparisons ($p > .05$) are not reported in full

Table 5. Summary of changes pre and post ACA - significance (p) values

OEFO Category	Measure	Station A[#]	Station B	Station C	Station D
Responsiveness	Turnout times	<.05 ↑	n/s	n/s	n/s
	Att. Times LCI	n/s	n/s	<.05↓	n/s
	Att. Times nLCI	n/a	n/a	n/a	n/a
Availability	WT pump	n/s	<.05↓	n/s	n/s
	RT pump	n/a	<.05↓	=.002 ↓	<.002 ↓
	Sickness	<.05↑	n/s	n/s	<.05 ↑
Preparedness	Drills per person	n/a	n/s	n/s	<.05 ↑
	HFSC	n/a	n/s	n/s	n/s

↑ improved performance ↓ reduced performance
n/s refers to not significant *p* >.05
n/a data not suitable for statistical analysis
Station A data compared to Counties means as pre DCP unavailable
LCI Life Critical Incident

4.2 Responsiveness

For life critical incidents (Table 6), Station C has remained stable for turnout times from pre- to post-ACA. Station A appears in line with the Counties mean for wholetime stations. The average turn-out time for Station B and Station D has increased (i.e. was slower) by 9% for both stations. Comparisons were non-significant for all stations; however, the comparison between Station A and the Counties means only failed to reach significance once the Bonferroni correction was applied (t (18.61) = -2.19, p = .042).

Attendance times (time mobile to time at scene) for life critical incidents have a more mixed picture. Station D has a reduced average attendance time, Station C has increased, and Station B has remained stable. Station A has a slower than average attendance when compared to the service average. The percentage of calls achieved within the Counties target time of under 10 minutes for life critical incidents (The Counties, 2012) remained high and showed a slight improvement for Station B and Station C. Station A shows a higher percentage of life critical incidents attended within the same target time when compared to the Counties mean. However, Station D shows a moderate decline in the percentage of life critical incidents attended within 10 minutes. The comparisons were again non-significant for all stations; however, Station C only failed to reach significance once the Bonferroni correction was applied (*t* (11) = -2.42, *p* = .034).

Table 6. Life critical incidents (turn-out time, attendance time)

	Mean № Life critical calls per month		Mean turn-out time (sec) (SD)			Mean attendance time (sec) (SD)				
	Pre DCP	Post DCP	Pre DCP	Post DCP	% Change	Pre DCP	PRE DCP % on target	Post DCP	Post DCP % on target	% Change
Counties mean	3.81		151.75 (4.94)			320.33 (17.33)	95.92	-	-	
Station A#	N/A	3.17	N/A	145.92 (7.80)	-2%*	N/A	N/A	360.17 (76.42)	98.33	+15%
Station B	3.5	3.41	160.33 (18.68)	175.58 (24.94)	+9%	431.83 (50.05)	92.75	453.75 (58.19)	93.08	+5%
Station C	4.25	3.91	179.42 (13.82)	186.17 (17.78)	+2%	351.75 (32.64)	94.00	378.25 (25.78)	96.75	+7%*
Station D	2.25	2.75	164.25 (23.11)	179.75 (13.54)	+9%	505.33 (74.92)	82.83	461.92 (59.87)	75.25	-6%

Comparison made against Counties mean
* significant at <.05 but non-significant after Bonferroni applied

The responsiveness data for attendance time to non-life critical incidents (normal road speed calls; Table 7) indicated that percentage attendance within the Counties target of 20 minutes (The Counties, 2012) has remained very high with a minimum of 98.92% pre-ACA and 98.83% post-ACA. ACA seems to have had little effect on attendance times for these types of incidents. The narrow variability in the data suggests this measure lacks the sensitivity to uncover any influence. The data was not scrutinized further as it is unlikely to yield any meaningful analysis.

Table 7. Non-life risk responsiveness (% under 20 min target time)

	Mean № non-life risk calls Pre-DCP per month	% attended within target time	Mean № non-life risk calls Post DCP	% attended within target time	% change
Counties mean	29.9	99.42	-	-	
Station A#	n/a	n/a	23.17	98.83	+1%
Station B	28.92	99.75	28.75	99.42	0%
Station C	35.00	98.92	30.09	99.73	-1%
Station D	16.33	99.00	17.33	99.50	-1%

comparison made against Counties mean

4.3 Availability

A summary of the availability data can be found in Table 8. The availability of the wholetime appliance was very high in both periods (98.56% or above) at all stations. There is a small reduction at all stations between pre- and post-ACA however this is small and potentially due to natural fluctuation. A more significant reduction is seen in retained appliance availability at the stations under investigation and across the wider service. This is most marked at Station D. Comparisons for wholetime availability for all stations were non-significant; however, Station B only failed to reach significance once the Bonferroni correction was applied (t (11) = 2.27, p = .045).

Table 8. Availability (wholetime and retained appliance, and sickness absence)

	Wholetime Pump available (%)			**Retained Pump available (%)**			**Sickness absence (days per person per year)**		
	Pre DCP	Post DCP	% Change	Pre DCP	post DCP	% Change	Pre DCP	Post DCP	% Change
Counties mean	99.45	99.24	-.21%	92.74	89.21	-4%	4.40	3.68	+16%
Station A#	n/a	99.28	-.17%	n/a	n/a	n/a		2.23	+49%
Station B	99.75	98.56	-1.19%*	97.69	95.15	-3%	3.58	4.34	-21%
Station C	99.41	98.97	-.44%	98.51	95.08	-3%	3.43	2.11	+38%
Station D	99.71	99.63	-.08%	93.49	68.25	-27%	7.18	2.98	+58%

comparison made against Counties mean
* significant at <.05 but non-significant after Bonferroni applied
** Significant at corrected p value .002

Retained availability was not significantly different between the two-time frames at Station B once the Bonferroni correction was applied (t (11) = 3.84, p = .003), however this is approaching the corrected p value. Retained availability was significantly reduced in the 12-months after ACA for Station C (Mean pre-ACA = 98.51, Mean post-ACA = 95.08, t (11) = 4.04, p = .002, r = .773), Station D (Mean pre-ACA = 93.49, Mean post-ACA = 63.25, t (11) = 10.64, p < .001, r = .955) and the Counties overall retained complement (Mean pre-ACA = 92.74, Mean post-ACA = 89.21, t (11) = 8.97, p <.001, r = .938). Large effect sizes were seen for all comparisons ($r \geq .5$). No retained appliance operates at Station A.

There is a large percentage reduction in sickness absence for Station C and D. Station B saw a large percentage increase in the same measure. These changes were not statistically significant in the 12-months before and after ACA for Station B and C. Station D failed to reach significance once the Bonferroni correction was applied (t (11) = 2.34, p = .039). There was a large percentage difference in sickness absence between Station A 12-months post ACA and the Counties means 12-months prior to ACA; however, this failed to reach significance once the Bonferroni correction was applied (t (16.94) = -3.22, p = .005) although this is approaching corrected p value.

4.4 Preparedness

The yield of data for measures of critical incidents, commendations and personal injuries were not high enough for meaningful comparisons, and so were omitted from the analysis. Data for equipment testing has very little variability due to being a routine action incorporated into ACA activities, and so was also omitted from the analysis. Equivalent before and after comparisons for hydrant checks are not possible due to increasing numbers being performed by non-operational personnel.

The greatest variability was found within data for training drills and Home Fire Safety Checks (HFSC). A summary of the data is found in Table 9.

Table 9. Preparedness (training drills and HFSC's)

	Mean № drills per station per month			**Mean № drills per person per month**			**Mean № HFSC's per month**		
	Pre DCP	Post DCP	% Change	Pre DCP	post DCP	% Change	Pre DCP	Post DCP	% Change
Counties mean	35.96						36.68		
Station A#	n/a	13.17	-63%	n/a	11.29		n/a	27.67	-25%
Station B	31.67	16.92	-47%	13.57	14.50	+7%	25.00	25.83	+3%
Station C	35.67	14.08	-61%	15.29	12.07	-21%	34.17	33.25	-3%
Station D	28.83	20.58	-29%	12.36	17.64	+43%	7.33	4.58	-38%

comparison made against Counties mean

The number of drills performed by the station appears to have had a large decline but this does not account for fewer personnel on station following the introduction of ACA. To control for the number of firefighters the mean number of drills per person per year were calculated, to allow for a meaningful comparison between time periods. Once this control measure was applied, the number of drills carried out before and after ACA remained relatively stable. Station D is an exception, where personnel carried out more drills per year in the 12-months following ACA introduction. None of the comparisons within this measure reached statistical significance; however, training drill per person at Station D was only failed to reach significance once the correction was applied (t (11) = -3.02, p = .012).

5 Discussion

The aim of this paper was to present a framework to measure operational effectiveness in the FRS. Utilising established approaches, the OEFO organised existing performance data in a meaningful way to allow comparisons pre- and post-ACA. The relative infrequency of some FRS activities, e.g. health and safety events, and insensitivity of some measures required data to be drawn from a variety of sources.

Overall, ACA was found to have mixed effects on operational effectiveness but no major positive or negative effects at any of the stations under investigation. Some areas may be more sensitive to this new working arrangement than others, such as retained appliance availability. The FRS should be aware of how these may impact on service delivery before implementing ACA into further stations.

5.1 Responsiveness

Turnout times for life-critical calls for Station A and Station C have remained stable, whereas Station B and Station D have seen moderately slower times in the 12-months following ACA. These moderate differences were not found to be statistically significant; however, for Station

D the mean turn out time in the 12-months post ACA did fall above the Counties target of three minutes (The Counties, 2012[a]). With three minutes being a critical point for the Counties.

Findings for life-critical attendance times are more mixed. Two of the existing stations, B and C, have seen a moderate increase, but the other existing station, D, has seen a moderate decrease in attendance times. The large difference between Station A and the Counties mean can be attributed to the differences in station ground; the area of coverage offered by each station. The Counties mean includes city fire stations which have a much smaller coverage and attendance times at city stations are well within the 10-minutes and much quicker than rural stations, such as Station A. The yearly means for all stations falls below the Counties target of under 10 minutes (The Counties, 2012[b]). However, at Station D the percentage of calls attended within the target time dropped by 7%. Station D covers a very large rural area, and so other variables, including distance of the incident from the station and weather conditions reducing the speed of the appliance, have greater influence. Whilst these factors apply for the other three stations, Station D covers 154 square miles compared to the next largest at 107 square miles. Attendance times have the greatest potential of interference from extraneous variables and the idiosyncratic nature of the stations make assumptions about the influence of ACA difficult. As part of the data gathered for each incident, any records given for attendance slower than 10 minutes would be particularly useful and could be examined for factors that may be linked to ACA.

Attendance at non-life-risk (normal road speed) incidents within the target (20 minutes) remains very high, above 98%, before and after ACA across all stations. The narrow variability in this data limits this measure's usefulness in assessing responsiveness. The data provided by the Counties was in the form of the number of non-life-risk calls and the percentage attended within target per station per month. Scrutiny of actual times attended in seconds would allow for comparable analysis to the life-critical incidents and increase the sensitivity of the data; the target time of 20 minutes may mask a variability across stations and across time.

Response times (turnout and attendance) appear to provide a useful measure of responsiveness, but the findings support the relative insensitivity of this data from a lack of variability (Bateman *et al.*, 2016; Carvalho *et al.*, 2007; Jaldell, 2005). The low number of life-critical calls in the data set reduces variability and the statistical power that can be provided by measures of responsiveness. As such the findings support the need to assess additional performance measures to provide a fuller picture of operational effectiveness.

5.2 Availability

Availability has had the most marked change following the introduction of ACA. Wholetime appliance availability has remained high across all stations. The strategic importance of these appliances may motivate the service to move resources, including firefighters, to keep appliances available.

Retained appliance availability has seen a large percentage reduction across all the stations and the service as a whole. For two of the stations, C and D, and for the Counties' entire retained complement this reduction was statistically significant, and Station B was approaching the corrected significance value. Interviews revealed that there was a reduction hours ACA staff were able to offer at retained stations as part of dual-contract, and dual-contract personnel ended retained commitments once starting ACA. The wider impact on retained availability should be considered by the FRS when implementing ACA, including non-ACA stations with a reliance on dual-contract personnel.

Sickness has seen a large decrease across two stations, C and D, as well as a large percentage difference between Station A and the Counties mean. These differences failed to reach significance once the correction was applied, however, the low starting point may mask an effect. From the last available reports capturing sickness absence data across the UK, FRS personnel were shown to have a higher than average levels of sickness absence when compared to other workers (6.3 vs 4.9 days per person per year) (Black and Frost, 2011; Department for Communities and Local Government, 2010). The Counties sickness absence rate of 4.4 days per person per year compare favourably to the UK FRS as a whole, potentially due to a healthy worker effect following the introduction of mandatory health screening, fitness tests, and flu inoculations for operational personnel (as stated by Occupational Health Manager, and Health and Fitness Advisor for the Counties, 2013); thus, the large percentage change is still worthy of further investigation. Interviews revealed informal arrangements cover through swapping of shifts rather than a lower incidence of ill health. Station B saw a large percentage increase in sickness absence. Following interrogation of the sickness absence data supplied, this increase in the mean appears to be due to one period of long-term sickness by one crew member skewing the data. The low starting point of this data means one extended period of sickness may influence the overall results and therefore no conclusions can be drawn as to whether Station B follows a different trend to the others.

5.3 Preparedness

Following the pilot study, procedural changes in the wider service impacted on post ACA data. Increasing numbers of hydrant checks are performed by specific non-operational personnel, reducing the need for stations to undertake these checks. Any difference between the two time-points cannot be attributed to ACA because of this procedural change.

The number of drills performed per person per month may be influenced by the introduction of ACA because the 12 hours down-time reduces the time available for training. Station C saw a large percentage decrease in the number of drills performed. The Bonferroni correction applied to the analysis may have led to a type II error for this variable, due to the conservative nature of the correction, thus reducing the power to uncover an effect (Gelman *et al.*, 2012). Interrogation of the raw emergency call data could uncover whether the time of day that Station C was called out has impacted on time available for training, through pushing back the start of the working day. By contrast, Station D has seen a large percentage increase in the number of drills performed following the introduction of ACA, which fell below the .05 significance level but failed to reach significance when the Bonferroni correction was applied. The low call profile for this station may be a factor in creating a greater proportion of time for training. Interviews also revealed the self-rostering nature of the shifts required risk critical training to be repeated more often to ensure all personnel maintain competency, potentially increasing the number of drills per person.

Training data was only available via station records, not for centralised training from the Training and Development Department. To maintain competency in key areas, breathing apparatus and live fire training are carried out centrally, involving the fire appliance from the station attending a training site with all the crew available that day. When using a watch-based system this would entail visiting centralised training together as one watch with all crew trained simultaneously. With the self-rostering on ACA the same personnel are rarely work together on a continuous basis. Training department data could be useful in quantifying whether this had any impact on the efficiency of the ACA stations through repeated visits to centralised training.

HFSCs are carried out in response to demographic information via the MOSAIC profiling system (Local Government Association, 2012) to identify those most vulnerable to the risk of fire. Stations are responsible for visiting those households to offer advice and fit smoke alarms where necessary. Stations B and C saw a minimal change in the number of checks performed pre- and post-ACA. One explanation for this could be that HFSCs are a targeted indicator, so this work may be prioritised over non-target driven work, such as training. Station D had a large percentage decrease in the number of HFSCs performed, however, a low starting point means that in real terms this difference is minor. The demographics of the Station D area may explain the lower figure of HFSCs when compared to the other stations in the analysis. There are much lower levels of social deprivation and higher than average employment levels particularly in professional and managerial occupations (District Council Statistics[2], 2015). As there are fewer vulnerable members of the community within this station area, this data will be more sensitive to change.

5.4 Strengths and Limitations of EOFO

Previously, there has been no established way to measure operational effectiveness within the FRS, along with disjointed methods of assessing performance. We have outlined the OEFO as a methodological approach for empirically measuring operational effectiveness. In addition, the OEFO offers a tool for meaningful comparisons between and within services as part of a structured performance management regime, such as that posited by the home secretary (May, 2016) and within the new inspectorate framework (Home Office, 2017[b]).

As this study makes use of existing performance data already collected by the Counties' FRS we suggest data already collected by Fire and Rescue Services has an excellent fit within the framework and adequately assesses operational effectiveness within this context. This minimises any impact on resources through the need for new data collection.

The Bonferroni correction was applied in order to minimise the risk of familywise error, but with this comes the risk of a type II error. The conservative nature of this test can lead to the retaining of the null hypothesis and severely reduce the power to detect an important effect (Gelman *et al.*, 2012). By evaluating the results from the statistical analysis alongside the

[2] Reference anonymised for maintenance of The Counties FRS anonymity

guidance and targets set by the Counties FRS within the IRMP and other performance documents, the risk of dismissing a change of importance to stakeholders is minimised. Stakeholder assessments of the practical significance of any differences should be considered alongside the results from the statistical analysis: non-significant differences may still be important to stakeholders.

A deeper exploration of the softer performance measures, such as qualitative descriptions behind attendance times, could provide a history behind the numbers. This could uncover the influence of any confounding variables not considered as part of the analysis, such as procedural changes or new technology. Qualitative exploration can also go deeper into reasons for any impact, as quantitative assessment such as this provides an answer to whether there has been any impact but fails to uncover why or how.

6 Conclusion

The development of OEFO provides a rigorous approach for the tracking of changes to fire stations over time. The approach reflects the societal requirements for good response times but takes in the need for crews to operate effectively (and safely) whilst ensuring that preventative activities central to public safety are also maintained. Our research considered how these aspects can be exemplified through the three dimensions of responsiveness, availability, and preparedness. Whilst it is desirable to express these in a single measure this is likely to be over crude and mask potential problems. The assessment of public services beyond cost and simplistic single measures is crucial to service delivery. It allows informed decision making and the monitoring of changes. It also ensures decision makers are held accountable and know that they will be held accountable by highlighting appropriate measures.

Using the OEFO tool, ACA was not found to have any overall positive or negative effect at the four stations under investigation. The purpose of this project was to confirm whether ACA, as an intervention to reduce operational cost without adversely affecting operational performance, met its objectives. ACA allows a lowering of total full-time fire fighters and so the aim of no loss of cover is largely substantiated; however, retained numbers may be affected indicting the service may lose depth of coverage. For day to day incidents there is no loss of coverage but major events that require many fire crews over a sustained period may be affected.

Retained availability needs further investigation to uncover the underlying mechanisms of negative results across the stations and the Counties as a whole, as this may influence the service's ability to respond to a major incident. Wholetime operational effectiveness has not seen a negative impact overall, and locally there have been improvements in sickness absence and training. Due to the flexibility afforded to the Watch Managers in the running of individual stations, local variations in how the system is managed could explain these improvements.

Effective performance centres on what is measured, and current systems of benchmarking performance within the FRS do not provide clear comparative elements. The OEFO provides a base upon which empirical work can be formed. The OEFO is the first tool to attempt to operationalize operational effectiveness within the FRS using an established and well-researched aggregate measure from operations management. Adapting OEE for use outside of manufacturing has previously produced a robust and relevant measure for the sector to which it is applied (Simons *et al.*, 2004), supporting the feasibility of the OEFO framework.

References

Bateman, N. Maher, K., and Randall, R. (2016), Drivers of change within the UK Fire and Rescue Service: an operations management perspective. In Z. Radnor, D. Upton, M. Kumar, A. Esain, S. Williams and N. Bateman (Eds) *Public Service Operations Management.* Routledge, Abingdon, Oxfordshire, pp. 138-155

Black, Dame C,. and Frost, D. (2011), *Health at Work: An independent review of sickness absence.* HMSO, London.

Carvalho, J., Fernandes, M., Lambert, V., and Lapsley, I. (2006), Measuring Fire Service Performance: A Comparative Study. *International Journal of Public Sector Management*, Vol 19 No 2, pp. 165-179.

Chief Fire Officer Association (2012), Operational Assessment and Peer Challenge Toolkit. CFOA, Tamworth, available at https://www.ddfire.gov.uk/sites/default/files/attachments/OpA_and_Fire_Peer_Challenge_Toolkit.pdf (accessed 23rd March 2018)

Department for Communities and Local Government (2010), *Fire and Rescue Service: Operational Statistics Bulletin (2009/10).* DCLG, London, available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/11454/1693248.pdf (accessed 23rd March 2018)

Department for Communities and Local Government (2012[a]), *Firefighters' Pension Scheme: Proposed Final Agreement.* DCLG, London.

Department for Communities and Local Government (2012[b]), *National Framework for Fire and Rescue Services.* DCLG, London

Department for Communities and Local Government (2013[a]), 'Fire Statistics Monitor: April 2012-March 2013; Available at https://www.gov.uk/government/statistics/fire-statistics-monitor-april-2012-to-march-2013 (accessed 23rd March 2018)

Department for Trade and Industry (2000), *Quality Cost Delivery: Seven measures for improved competitiveness in manufacturing industry.* URN00/754. DTI/Pub4868/3k/9/00/NP. Department of Trade and Industry, London.

[District Council] (2015), *Key Statistical Data.* District Council, Economic Development.

Downe, J., Martin, S., and Doering, H. (2017), Peer Challenge: A sector-led approach to performance improvement in Fire and Rescue Services. In Murphy, P and Greenhalgh, K. (Eds) *Fire and Rescue Services: Leadership and Management Perspectives.* Springer, London, pp. 61-75.

Fire Brigades Union (2017), Dramatic increase in fire deaths as firefighter job cuts worsen. Available at: https://www.fbu.org.uk/news/2016/08/18/dramatic-increase-fire-deaths-firefighter-job-cuts-worsen (accessed 12th February 2018)

Fire Services Act (2004), Available at: https://www.legislation.gov.uk/ukppa/2004/21/pdfs/ukpga_20040021_en.pdf (accessed 20[th] October 2017)

Gelman, A., Hill, J., and Yajima, M. (2012), Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness.* Vol. 5, pp.189-211.

Home Office (2017[a]), Fire and rescue incident statistics: England, year ending September 2017. Available at:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/679924/fire-and-rescue-incident-sep17-hosb0418.pdf (accessed 12th February 2018)

Home Office (2017[b]), Fire and Rescue Framework for England: Government consultation document. Available at: https://www.gov.uk/government/publications?departments%5B%5D=home-office&publication_filter_option=consultations (accessed 7[th] January 2018)

Home Office (2018), Fire statistics (Workforce and Workforce Diversity). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/696917/fire-statistics-data-tables-fire1101.xlsx (accessed 17th April 2018)

Jaldell, H. (2005), Output specification and performance measurement in fire services: An ordinal output variable approach. *European Journal of Operational Research.* Vol. 161, pp. 525-535.

Kloot, L. (2008), Performance measurement and accountability in an Australian fire service. *International Journal of Public Sector Management.* Vol. 22, No. 2pp. 128-145.

Knight K. (2013), Facing the Future: *Findings from the review of efficiencies and operations in fire and rescue authorities in England* DCLG, London, ISBN 978-1-4098-3887-6

Local Government Association (2012), [*The Counties] Fire and Rescue Service Fire Peer Challenge: Final Report.* Local Government Association, London.

Local Government Association (2015), *Operational Assessment and Fire Peer Challenge.* Local Government Association, London.

May, T. (2016), Home Secretary Speech on Fire Reform. Delivered on 24[th] May 2016. London, Reform. Available at: www.gov.uk/government/speeches/home-secretary-speech-on-fire-reform (accessed on 20[th] October 2017)

Muchiri, P. and Pintelon, L. (2008), Performance measurement using overall equipment effectiveness (OEE): Literature review and practical application discussion. *International Journal of Production Research.* Vol. 46 No. 13, pp. 3517-3535.

Murphy, P. and Greenhalgh, K. (2013), Performance Management in fire and rescue services. *Public Money and Management.* Vol. 33No. 3pp. 225-232.

Nakajima, S. (1984), *Introduction to TPM: Total Productive Maintenance.* Japan Institute for Plant Maintenance, Tokyo, Japan (English Translation: Productivity Press 1988)

Neely, A., (1999), The performance measurement revolution: why now and what next?. *International journal of operations & production management*. Vol.19, No. 2, pp. 205-228.

OECD. (2004), Public Sector Modernisation: Governing for Performance [Online]. OECD, Paris.

Porter, M. (1996), What is Strategy? In Mazzucato, M. (Ed) *Strategy for Business: A Reader*. Sage Publications Ltd., London.

Radnor, Z., Osborne, S., and Glennon, R. (2016), "Public Management Theory" in Ansell, C. and Torfing, J. *Handbook on Theories of Governance*. Edward Elgar Publishing Ltd., Cheltenham

Simons, D., Mason, R., and Gardner, B. (2004), Overall vehicle effectiveness. *International Journal of Logistics Research and Applications.* Vol. 7, No. 2pp. 119-135.

Slack, N., Chambers, S. and Johnston, R. (2010), *Operations Management. 6th Ed.* Pearson Education Ltd, Harlow.

Taylor, B. (2016), Spatial modelling of emergency service response data. *Journal of the Royal Statistics Society: Series A (Statistics in Society).* 10.1111/rssa.12192

The Counties FRS (2012[a]), Policy Committee Report, 9th May 2012. [The Counties FRS]

The Counties FRS (2012[b]), Integrated Risk Management Plan 2013-2016. [The Counties FRS]

The Counties FRS (2012[c]), Policy Committee Report. [The Counties FRS]

The Guardian (2017), Fire deaths rise by 21% as chiefs' issue cuts warning. Available at https://www.theguardian.com/uk-news/2016/apr/26/fire-deaths-rise-by-21-as-chiefs-issue-cuts-warning (accessed 12th February 2017)