bureau of health information

Spotlight on Measurement

Measuring and reporting performance of NSW ambulance services

BUREAU OF HEALTH INFORMATION

Level 11, 67 Albert Avenue
Chatswood NSW 2067
Australia
Telephone: +61 2 9464 4444
Email: BHI-enq@health.nsw.gov.au
bhi.nsw.gov.au

State Health Publication Number: (BHI) 170323
ISSN: 2204-1598 (Print), 2204-552X (Online)

Suggested citation:
Bureau of Health Information. *Spotlight on Measurement: Measuring and reporting performance of NSW ambulance services.* Sydney (NSW): BHI, 2017.

Please note there is the potential for minor revisions of data in this report.
Please check the online version at **bhi.nsw.gov.au** for any amendments.

Published June 2017

Table of contents

Foreword

From June 2017, the Bureau of Health Information (BHI) is incorporating information on the activity and performance of ambulance services in NSW into its quarterly reporting products.

At first glance, it may seem that reporting on ambulance performance is straightforward – measurement of the time taken from a triple zero call to a patient's arrival at hospital appears to be easily achieved with available information systems. However, meaningful measurement and assessment of performance is more complex than that. Patients' healthcare needs vary in urgency; ambulances are based in geographic regions of different size and hospital density; and as a provider of mobile emergency healthcare, the ambulance service is subject to unanticipated events and demand pressures at a state, regional and local level.

This edition of *Spotlight on Measurement* details the steps taken in light of this complexity to develop performance measures suitable for public reporting. It describes the options considered and analyses used to guide decisions about what to measure, how to compare performance, what to report and how to guide interpretation of results.

The introduction of quarterly releases of ambulance data has brought with it several advances in performance measurement and reporting in NSW.

First, it provides new insights into previously unreported aspects of performance within the multifaceted NSW healthcare system. It contrasts system and patient perspectives – comparing measures of the time it takes ambulance teams to reach patients, with measures of the time patients wait from calling for assistance until help arrives.

Second, it catalysed a clear step forward in performance measurement and reporting methods. The work has presented an opportunity to develop a way to assess system resilience in the event of unexpected surges in the demand for services. This new approach clearly has potential applications in other healthcare contexts, particularly where emergency responses are an integral part of service provision.

Third, it opens the way to help understand the how different parts of the healthcare system interact, particularly in emergency responses across community and hospital settings.

The project brought together experts on mobile emergency healthcare services, statistical analysis, research design, public reporting and the communication of complex healthcare information. Their efforts have resulted in an integrated approach that provides accountability and informs efforts to improve care locally.

Kim Sutherland
Acting Chief Executive,
Bureau of Health Information

10 key points

1. **Responses will be the core unit of measurement for analysis and reporting** – defined as a dispatch of an ambulance following a triple zero call.

2. **Measurement and reporting will reflect quarterly activity and performance, with a particular focus on timeliness** – analyses show quarterly results and changes over time, taking into account seasonality.

3. **Quarterly reporting will be established with four key measures** – each reflects different time intervals of the ambulance response and together, the measures provide both operational and patient perspectives. The measures are: mobilisation time, turnaround time, response time and call to ambulance arrival time.

4. **Different statistics are used to reflect the key measures** – median and 90th percentile for more operational measures, and percentage within a defined time period for patient perspective measures.

5. **Assessments of timeliness will be reported using existing priority definitions of emergency (priority 1) and urgent (priority 2) categories.** A third priority category (priority 3, time-critical) was considered for reporting but included 'booked calls', that confounded timeliness assessment.

6. **Activity and performance will be reported at a NSW and zone level.** Levels of random variation and non-modifiable factors (such as distance and travel time) limit the ability to report local response area results on a nominal (named) basis.

7. **The number and distribution of local response areas outside performance thresholds will be reported.** This will allow tracking of changes over time. The threshold will be met if over 90% of call to ambulance arrival times within a local response area are within 30 minutes for priority 1 responses, and within 60 minutes for priority 2 responses.

8. **Additional contextual information will be used to make results meaningful and assessments fair.** This includes descriptive information (highlighting changes in priority code definitions or in policies regarding transportation of patients), and analytic approaches (stratification of results by 'local response area types' – clustered on the basis of size and staffing of local response areas).

9. **Activity and performance will be considered in an integrated way** – both in terms of the interface and between ambulance and emergency department services and in terms of relative performance on days of high activity and days of low activity that differ significantly from usual patterns of demand (surge and lull days).

10. **Further development in ambulance performance measurement is planned.** This will focus on clinical measures of effectiveness and appropriateness of care and measures of integration of care across ambulance and hospital sectors.

Summary

This report outlines the steps taken to develop measures of activity and performance for public reporting about ambulance services in NSW. Focusing on measures of timeliness, it describes the information and analyses that guided decisions about which aspects of performance to measure, the metrics to report, and how to ensure fair assessment and meaningful attribution.

The measures will form the basis for a new module of BHI's *Hospital Quarterly* report, renamed *Healthcare Quarterly*. The reporting seeks to strike a balance between including sufficient measures to provide a rounded picture of performance, but not so many as to overwhelm interpretation and action.

What to measure

Ambulance-based electronic information systems collect a range of data points that offer several options for measurement. The indicator development process considered the merits of different bases for measurement – such as calls, incidents or responses – and canvassed reporting regimes in other healthcare systems.

Responses, which refer to the dispatch of an ambulance, were selected as the unit of measurement for assessing performance for two key reasons. First, responses reflect the demand for and utilisation of ambulance vehicles and are salient and meaningful to a range of audiences. Second, they are widely used in NSW by the ambulance service for operational purposes and elsewhere in other systems as the primary unit of timeliness metrics.

Capturing events of interest

Different outcomes or events of interest can provide insight into various perspectives on performance. For ambulance services, measures should be informative operationally and reflect efficiency and preparedness of services. There is also a clear imperative to consider performance from the patients' perspective.

Following assessment of various measures, a multi-metric approach to performance reporting will be used. It includes:

- Three operational measures which reflect different time intervals of the ambulance response. The measures are mobilisation time, response time and turnaround time.
- A patient-focused measure, the 'call to ambulance arrival time', which reflects the time patients wait from calling for assistance until an ambulance arrives.

In terms of reporting, median and 90th percentile times are proposed for operational measures. For the patient-focused measure, the percentage of call to ambulance arrival times within defined times (for different priority categories) are proposed.

Intra-zone variation in performance in call to ambulance arrival times is reported non-nominally as the number of local response areas for which 90% of priority 1 (P1) call to ambulance arrival times were within 30 minutes; and the number for which 90% of priority 2 (P2) call to ambulance arrival times were within 60 minutes.

Making fair comparisons

In developing performance measures, different subgroups based on geographical units or zones, organisational characteristics, and priority category were explored.

Priority (or urgency) category was shown to be an important factor to be considered in performance assessment.

Ambulance services in NSW are organised into more than 250 local response areas (LRAs). Each LRA is centred around an ambulance station. LRAs differ in their staffing and organisational arrangements and provide services across the spectrum of urgency categories.

Once acuity is taken into account however, LRA types did not explain a significant amount of unmodifiable performance. Therefore, the proposed approach for quarterly reporting is to stratify results based on priority category only. LRA type strata will be used to assess levels of intra-zone variation and for context only.

Performance attribution

Analyses examined the extent to which variation in different performance measures was attributable to LRAs, zones, or the system as a whole. Key to this assessment was understanding unmodifiable features in different geographical areas that could affect performance.

At the LRA level, results show that variation in performance is not always meaningful or actionable locally. Therefore *Hospital Quarterly* will not include nominal (named) reporting of LRAs.

Considering the statewide performance as a whole however, *Healthcare Quarterly* will describe the type and location of LRAs where patients wait beyond defined times, and monitor changes over time.

Nominal reporting will be used for zones

Suppression rules were established to ensure statistical stability of LRA results. Reporting thresholds were defined so that results unduly affected by small volumes are not reported publicly. LRA results will be shown on a non-nominal basis to illustrate intra-zone performance. For presenting these LRA results, those with fewer than five consecutive quarters of data, those with, on average, fewer than 100 responses per quarter, those with a coefficient of variation of over 10%, and the volunteer and community LRA type are suppressed. After the application of all rules, there were 113 LRAs with suppressed results.

Applying performance measures

As a provider of mobile emergency healthcare, the ambulance service can be affected by unanticipated fluctuations in demand. Resilience – the ability to respond to unexplained surges in demand, or to cope with unanticipated weather events, epidemics, or localised or statewide disasters – is an important characteristic of high-performing emergency services.

A statistical model was developed to enable categorisation of each day in a reporting quarter as either a day that was no different than expected in terms of activity, a day of unexpectedly high activity (surge day), or a day of unexpectedly low activity (lull day). Daily performance, across a suite of timeliness measures, was mapped against the observed number of responses on a given day, and the expected level of activity.

Turnaround time – the period between an ambulance's arrival at a hospital and the time it is 'clear' and ready to respond to another call – provides an opportunity to explore the interface between ambulance services and emergency departments. Using comparative data for LRA crews that take patients to multiple hospitals, and for hospitals that receive patients from multiple LRAs, median turnaround time results were more consistent with hospitals than with LRAs. This suggests hospitals have a greater influence on turnaround times than LRAs.

Moving forward

The introduction of public reporting of ambulance data represents a significant advance in transparency and accountability in NSW, and offers significant opportunities to improve care. Future work will focus on the development of more clinically-oriented measures that assess outcomes and processes of care and will supplement existing timeliness performance measures.

Setting the scene

About this report

The *Spotlight on Measurement* series focuses on methodological issues, describing the analytic approaches that were explored in indicator development and validation, and illustrating the results of key sensitivity analyses.

Background

NSW Ambulance provides frontline pre-hospital and out-of-hospital emergency patient care and transport services to over 7 million people across an area of 800,000 square kilometres in NSW.

Four control centres – located in Sydney, Charlestown, Warilla and Dubbo – receive calls, assign priority codes according to predefined criteria, and dispatch ambulances and crews.

Helicopter and fixed-wing transport and retrieval services are provided through a statewide specialist aeromedical control centre.

NSW contains eight sectors comprising 18 zones (11 in metropolitan areas of Sydney and seven in non-metropolitan NSW) (Figure 1).

Zones incorporate local response areas (LRAs) which are the smallest organisational units and are generally centred around an ambulance station.

There are four main types of LRAs, distinguished by staffing and organisational arrangements: 24-hour, 24-hour (with on call), non-24-hour, and volunteer and community (Figure 2). The distribution of types of LRAs varies across zones (Figure 3).

Figure 1 Local response areas across NSW

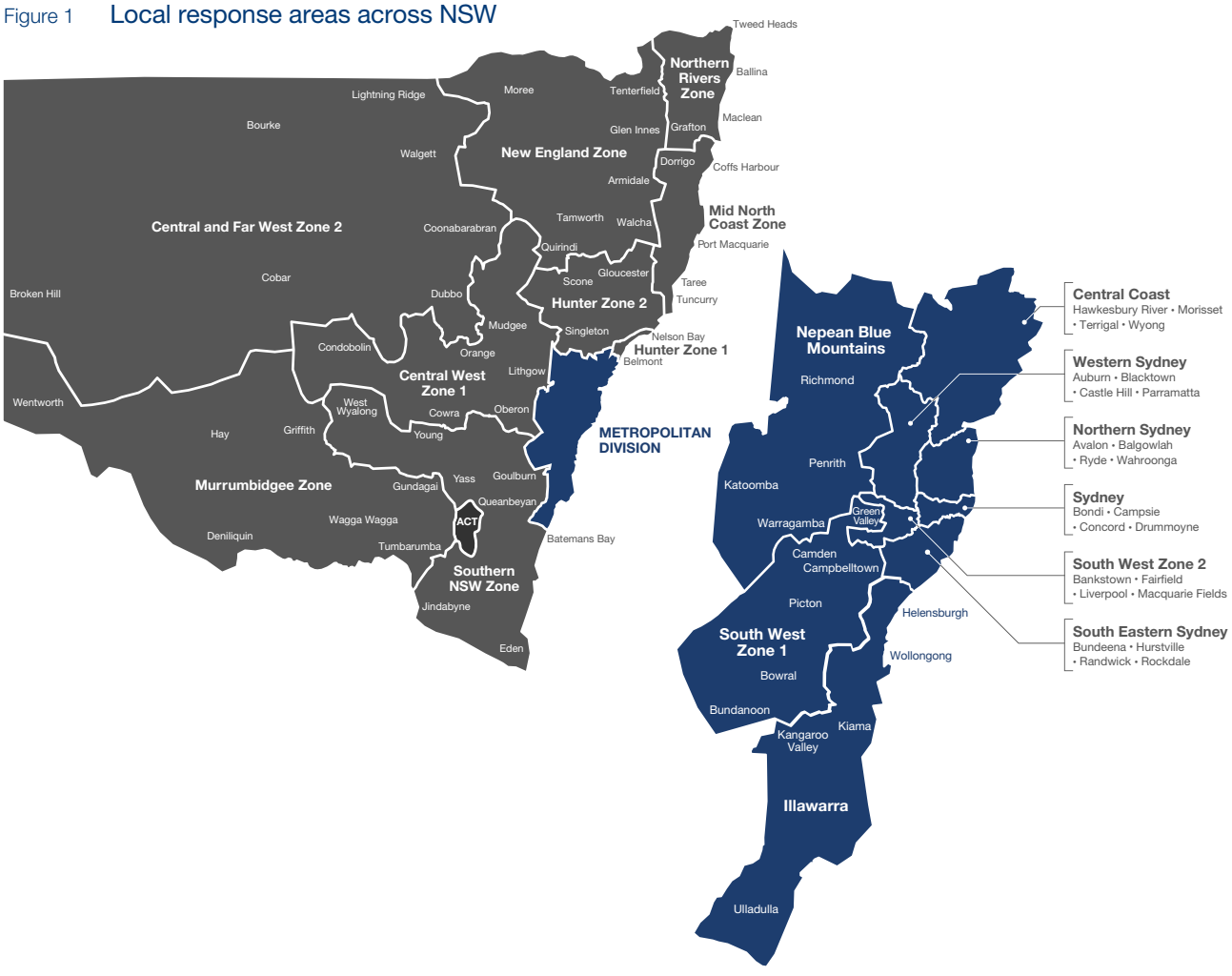
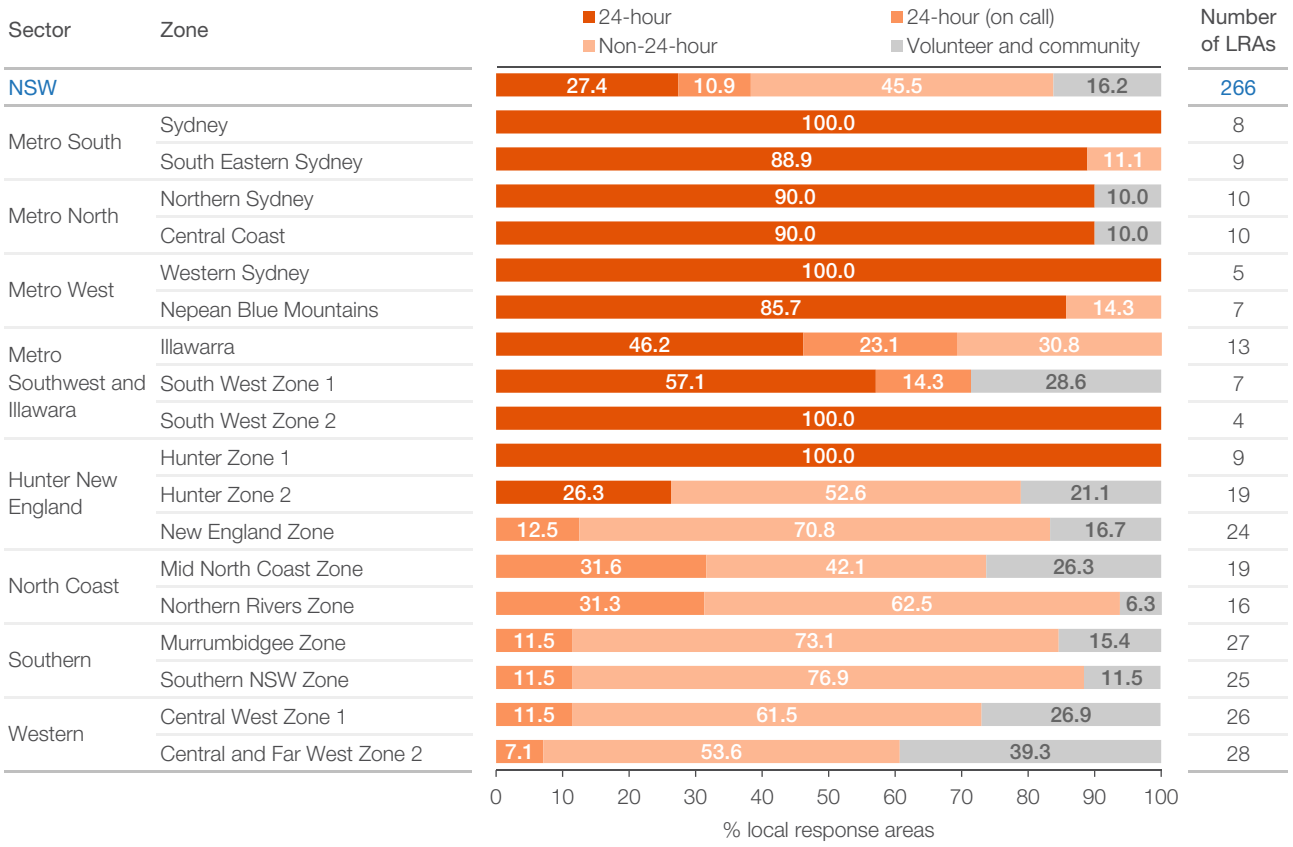


Figure 2 Organisation of ambulance services, NSW

Local response area type	Service type	Description
1	24-hour	Primarily situated in urban areas providing mostly urban, 24-hour operation. These are higher volume response areas, the majority with multiple vehicles and ambulance staff.
2	24-hour (with on-call)	Primarily situated in regional areas providing 24-hour operation, supplemented with on-call staff.
3	Non-24-hour	Primarily situated in regional and rural areas providing 8-, 12- or 16-hour operation with remaining time covered by on-call staff.
4	Volunteer and community (volunteer ambulance officers, community first responder programs and community-initiated groups)	Volunteer and community provide a first response and transport role in more remote areas. Some are attached to smaller stations, work with certified paramedics and respond in an ambulance vehicle. Members of community first responder programs are attached to emergency services, such as Fire and Rescue NSW, NSW Rural Fire Service and the NSW State Emergency Service, and respond in their agency vehicle. Community-initiated groups (not attached to a response agency) can form a community first responder unit. Members agree to be available on a regular basis and respond from within the community in a private or community-funded vehicle.

Figure 3 Distribution of local response area types, by zone, NSW, July to September 2016



Priority category and urgency

The ambulance dataset contains key variables that can be used to develop performance measures. The data describe the urgency of the incident, as assessed by operators taking the initial triple zero call.

Priority is classified into nine main categories – three of which (priority categories 1, 2 and 3) are useful for measurement of time-critical services and feature in this report (Figure 4).

There have been changes in recent years to the treatment of the priority category 3 (P3). In 2012, it was referred to as 'priority medical' and encompassed transporting patients to a health facility within an agreed time as determined by a medical practitioner or nurse for admission, including hospital to hospital transfers. Currently referred to as 'time-critical', priority 3 is a heterogeneous category that poses particular measurement challenges.

While the 24-hour LRAs deliver the majority of the most urgent priority 1 responses, the distribution of priority categories is similar across LRA types (Figures 5 and 6). Likewise, the proportion of responses in the various priority categories is similar across geographical zones (Figure 7).

Figure 4 Priority categories used to classify urgency of an incident, NSW

Code	Priority	Description	Example	Response required
1	**1A Emergency**	Highest priority – life-threatening case	Cardiac or respiratory arrest, unconscious, ineffective breathing	Immediate response – median within 10 minutes – under 'lights and sirens'
	1B Emergency	High priority	Unconscious	Emergency response – under 'lights and sirens'
	1C Emergency	Priority	Breathing problems, chest or neck injury, serious haemorrhage	Emergency response – under 'lights and sirens'
2	**Urgent**	Urgent	Abdominal pain	Urgent response without 'lights and sirens' within specified timeframes
3	**Time-critical**	Time-critical	Medical responses requested by medical practitioners, often pre-booked	Undelayed response within specified timeframes
4-9	**Non-emergency**	Non-emergency	Routine transport	Routine

Figure 5 Distribution of responses in priority category, by local response area type, NSW, July to September 2016

Priority category	Responses
P1: Emergency	127,986
P2: Urgent	121,250
P3: Time-critical	25,220
P4–9: Non-emergency	10,281

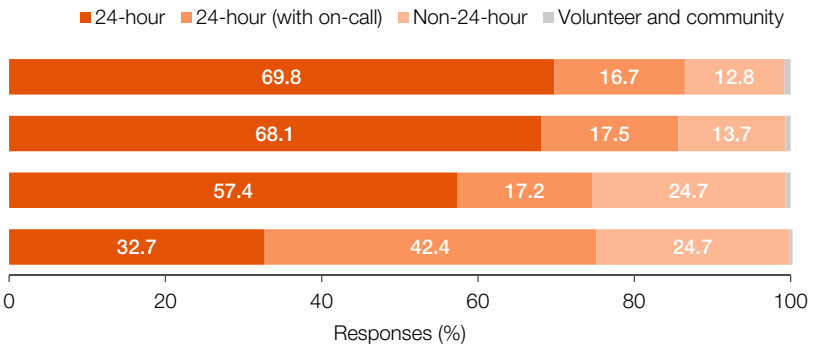


Figure 6 Distribution of responses in local response area types, by priority category, NSW, July to September 2016

Local response area	Responses
Peer group 1: 24-hour	189,730
Peer group 2: 24-hour (with on call)	51,292
Peer group 3: Non-24-hour	41,702
Peer group 4: Volunteer and community	2,013

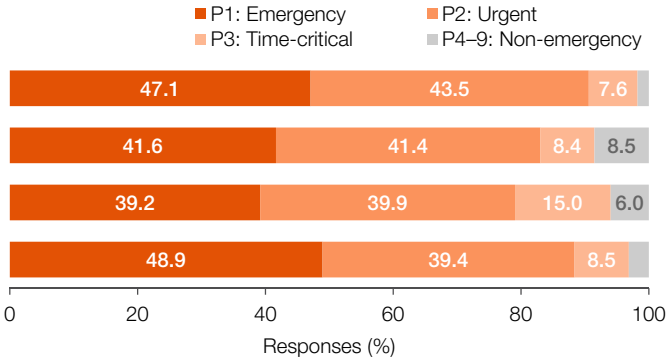
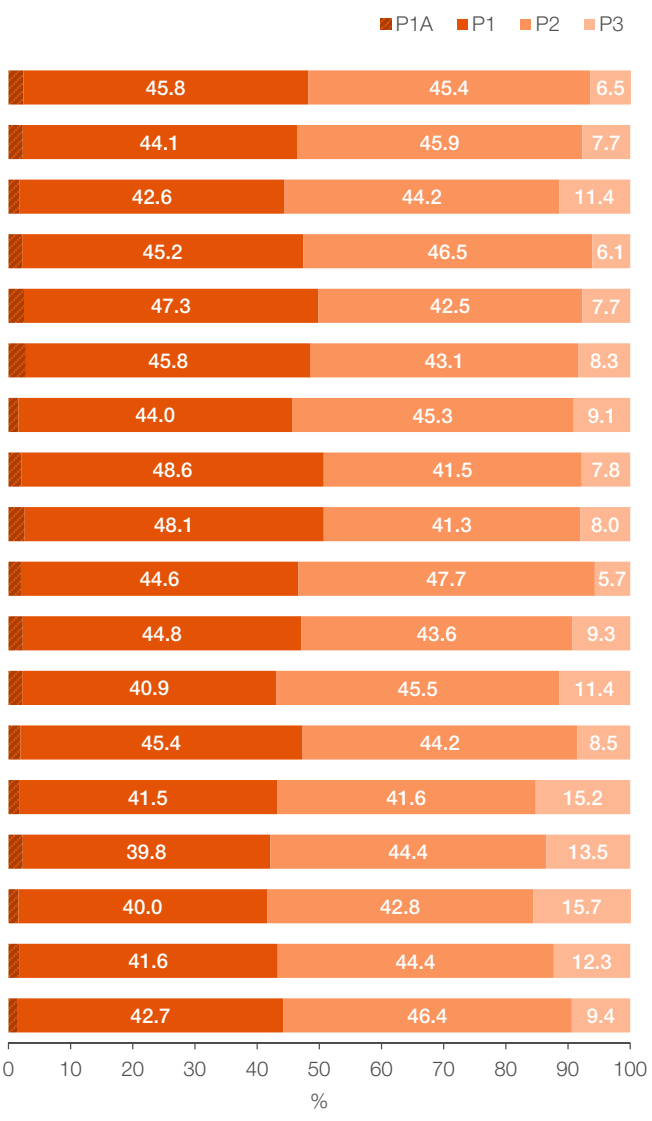


Figure 7 Distribution of responses in ambulance sectors and zones, by priority category, NSW, July to September 2016

Sector	Zone	Total number of responses P1–P3
Metro South	Sydney	26,767
	South Eastern Sydney	19,577
Metro North	Northern Sydney	20,364
	Central Coast	15,621
Metro West	Western Sydney	23,422
	Nepean Blue Mountains	16,241
Metro Southwest and Illawara	Illawarra	17,939
	South West Zone 1	14,596
	South West Zone 2	18,128
Hunter New England	Hunter Zone 1	15,137
	Hunter Zone 2	10,329
	New England Zone	8,230
North Coast	Mid North Coast Zone	14,704
	Northern Rivers Zone	14,135
Southern	Murrumbidgee Zone	11,553
	Southern NSW Zone	11,857
Western	Central West Zone 1	9,464
	Central and Far West Zone 2	6,392



Data and methods

Data source

The analyses in this report are based on NSW Ambulance data that include all incidents, responses and transports in NSW. Data were extracted by NSW Ambulance from the computer aided dispatch (CAD) system without any modifications. The CAD system captures the time stamps that are recorded electronically by operators and paramedics. Each quarter there is a (generally) small number of 'outages' where data capture is interrupted. NSW Ambulance estimates the number of responses that occurred during such outages. The data in this report exclude these estimates.

This edition of *Spotlight on Measurement* outlines the process of developing ambulance performance measures. Analyses were primarily based on the July to September 2016 quarter. Analyses for additional development and validation purposes used subsequent quarters of data. Statistical analyses were performed using SAS 9.3.

Data quality

Data quality checks focused on verifying the validity of underlying activity and time stamps. Invalid observations represented a very small percentage of the data and did not introduce a risk of bias in reporting. For example, in the July to September 2016 quarter, some exclusions based on data quality issues were:

- 14 incidents (out of 225,472) with a missing priority code
- 63 responses (out of 285,434) that had invalid response times (e.g. ambulance arrival on the scene preceded dispatch time).

Methods

Stratification analyses were conducted to delineate within and between organisational unit variation. Strata were defined on the basis of priority, geography and organisational characteristics. Various measurement statistics were considered for reporting operational and patient-focused measures (Figure 8).

Statistical modelling was used to establish patterns of 'normal variation' in the number of daily responses at a NSW level. The predicted number of responses for each day of the quarter was estimated from the model using a daily series spanning five years.

Figure 8 Measurement statistics

Measure	Definition
Median	As the middle value of a distribution, the median is a measure of central tendency that is robust to outliers. Medians are reported for mobilisation time, response time and turnaround time.
Mean	The mean is the average or the sum of all values divided by the number of values in the distribution. In non-normal distributions it can be disproportionately affected by a few extreme outliers.
Mode	A value that appears the most often in the distribution. The mode is the preferred statistic when the distribution is made up of discrete values and frequency is meaningful.
90th percentile	Often reported alongside the median, the 90th percentile provides insight into the tail of a distribution. For timeliness measures, it captures the time in which 90% of the sample received the element of care of interest. 90th percentiles are reported for mobilisation time, response time and turnaround time.
Percentage	Percentages can be used to reflect levels of compliance with recommended care processes or achievement of timeliness targets. Percentage of call to ambulance arrival times within defined times are reported.

All measures were validated by a two-stage process: internally via double coding of any analyses, and externally via feedback from NSW Ambulance for precision and sense checks.

Statistics

Statistics such as average or mean, median, percentile and percentage are options for descriptive reporting (Figure 8).

Ambulance response time data are not normally distributed, and a measure of central tendency that is robust to outliers is required (Figure 9). The long right-side tail indicates that the median is a more meaningful measure of central tendency than the mean (which is disproportionately influenced by outliers). The most extreme values (or the extent of the tail) is also important to assess, which is why medians are generally supplemented by 90th percentile values.

Interpreting data

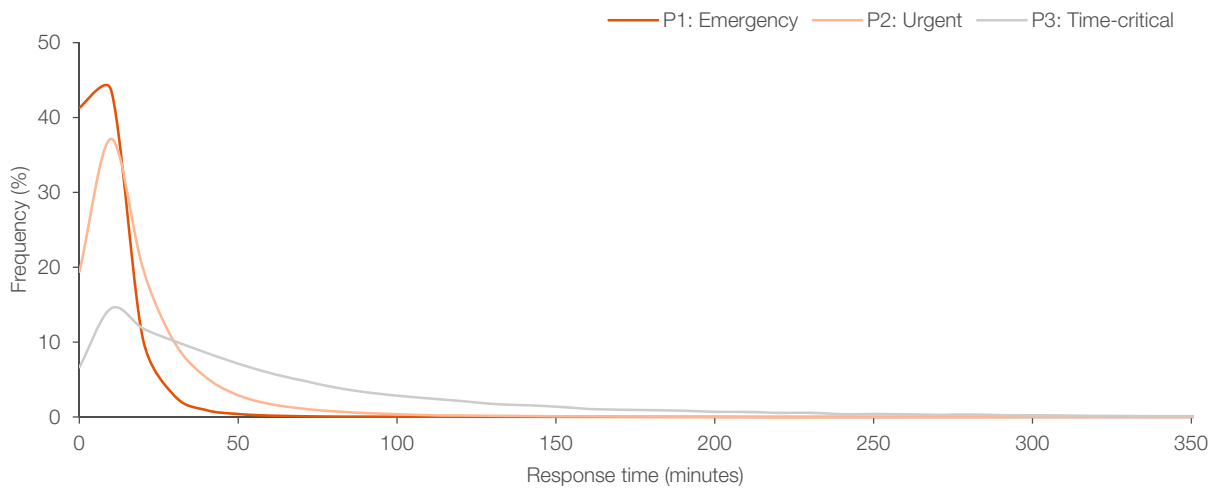
As with all statistics, interpretation of ambulance measures requires care, particularly in light of:

- A lack of clinical outcome data and limited ability to assess the effectiveness and appropriateness of care
- Potential unmeasured confounding factors such as imprecise priority classifications, or specific localised externalities such as traffic conditions or topography
- Difficulties parsing out and attributing performance in a multi-faceted health system and an ambulance system network
- A lack of a firm empirical or clinical basis for timeliness thresholds.

National and international comparisons

There are a number of different measures used to assess the performance of ambulance services. While measures appear similar across jurisdictions, there are significant definitional and operational differences that confound direct comparison and benchmarking. A summary of measures across different territories was compiled to inform the development of indicators for this report (Figure 10).

Figure 9 Distribution of response times by priority category, NSW, July to September 2016



Note: For the purposes of this graph, the x-axis is truncated at 350 minutes (for priority 1, observations extended to 270 minutes, for priority 2, 420 minutes and for priority 3, 1590 minutes).

Figure 10 Selected ambulance timeliness measures in different jurisdictions[2-10]

Place	Timeliness indicator
NSW Ambulance and Ministry of Health	Priority 1A median response time Priority 1 median response time Triple Zero grade of service Activation time within two minutes Mobilisation time within three minutes Case cycle time within 60 minutes
Queensland	Response times (min) median and 90th percentile by priority code Response time targets % <30 minutes: code 2B Response time targets % <60 minutes: code 2C
Victoria	Percentage of code 1 incidents responded to within 15 minutes (statewide target 85%) Percentage of code 1 incidents responded to within 15 minutes (urban central locality target 90%) Average response time for code 2 incidents
South Australia	Percentage priority 1 (life-threatening) ambulance intervention within 8 minutes Percentage priority 2 (potentially life-threatening) ambulance intervention within 16 minutes Percentage priority 3 (life-threatening) ambulance intervention within 30 minutes
Western Australia	Ramping: hours that ambulances wait outside ED when the hospital ED is unable to accept patients Off-stretcher target: percentage off stretcher in 20 minutes
Productivity Commission (Australia)	Median and 90th percentile (code 1 incidents) time taken between the initial receipt of the call for an emergency ambulance at the communications centre and the arrival of the first responding ambulance resource at the scene of an emergency
Canada	Average response time (mean = total calls/total time). Time from crew notification to arrival at scene. Legislated target is 90% in 24 minutes (Ontario) Percentage of responses within 8 minutes for severe cases requiring resuscitation and 6 minutes for a call requiring a defibrillator (Ontario) Percentage of responses less than 9 minutes for an ambulance to arrive on the scene in urban and metropolitan areas (British Columbia)
Wales	Category A median and mean response times by local health board by emergency level Category A percentage emergency responses arriving at the scene within 8 minutes (target is at least 65% at national level and 60% at local level) Percentage receiving call back for clinical care within 10 minutes Planned clinical telephone assessment: proportion receiving call back for clinical triage within 10 minutes Category C planned transport: time to face-to-face assessment – proportion of first responses within 30 minutes Category C planned clinical telephone assessment: proportion of calls receiving call back for clinical triage within 10 minutes
England	Time to answer call (seconds): median, 95th, 99th percentile Time to arrival of ambulance (minutes): median, 95th, 99th percentile Category A response time should be by 8 minutes 0 sec in 75% of cases Number and proportion of Red 1 calls resulting in an emergency response within 8 minutes 95th percentile of time from Call Connect of a Red 1 call to an emergency response arriving at the scene of the incident Number of Red 2 calls resulting in an emergency response Number and proportion of Red 2 calls resulting in an emergency response within 8 minutes Number of Category A calls resulting in an ambulance arriving at the scene of the incident Number and proportion of Category A calls resulting in an ambulance arriving at the scene within 19 minutes

Framework for developing measures

This edition of *Spotlight on Measurement* is organised according to BHI's framework for developing performance measures (Figure 11). It focuses on four key analytic questions:

1. What unit of ambulance activity should form the basis for performance assessment? Analyses considered options for 'cohorts' or units of analysis. Generally this is the denominator for measurement. For ambulance activity, potential units include calls, patients, vehicles, staff or responses.

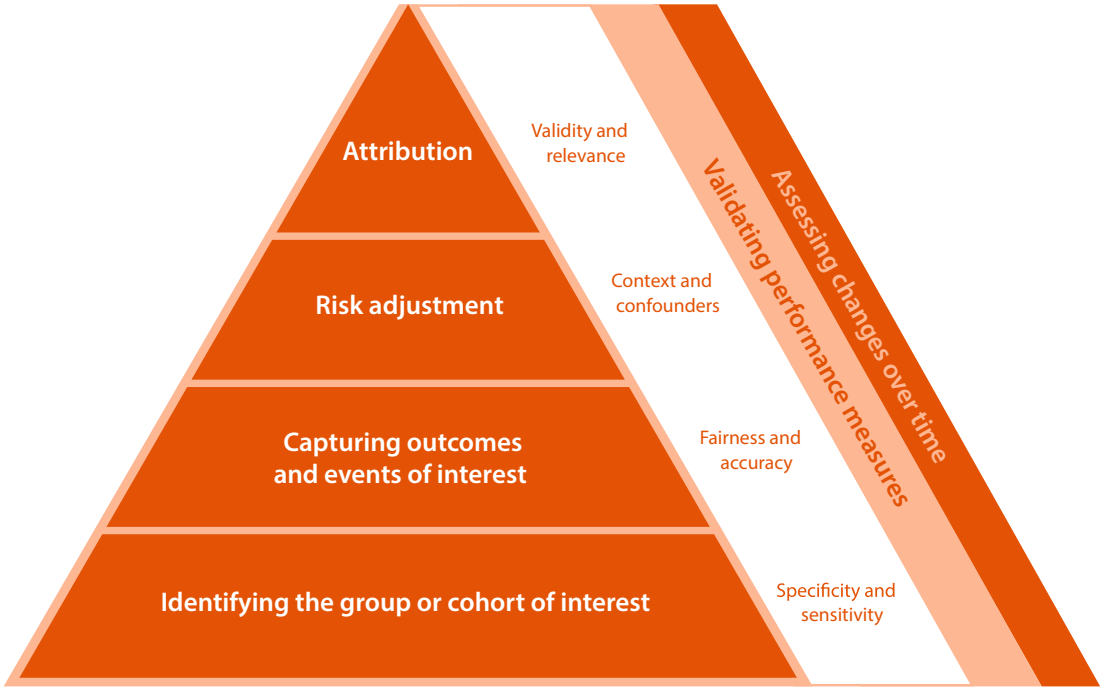
2. What aspects of performance or events of interest should be measured? Ambulance services can be considered in terms of a cycle. Measurement requires decisions to be made as to whether to focus on the entire journey or whether to highlight specific intervals to capture variation between units and over time.

3. How can fair comparisons be made? What factors can explain variation in performance? How should organisations and activities be clustered to ensure fair comparisons?

4. Where should performance be attributed? Critical to performance measurement is assessing whether variation is modifiable and, if so, what level of attribution can be made.

While measure development was based on empirical assessments, it was also informed by referencing other jurisdictions and by extensive stakeholder engagement with key organisations in NSW.

Figure 11 The Bureau of Health Information framework for performance measure development



Defining and validating measures

Defining the cohort

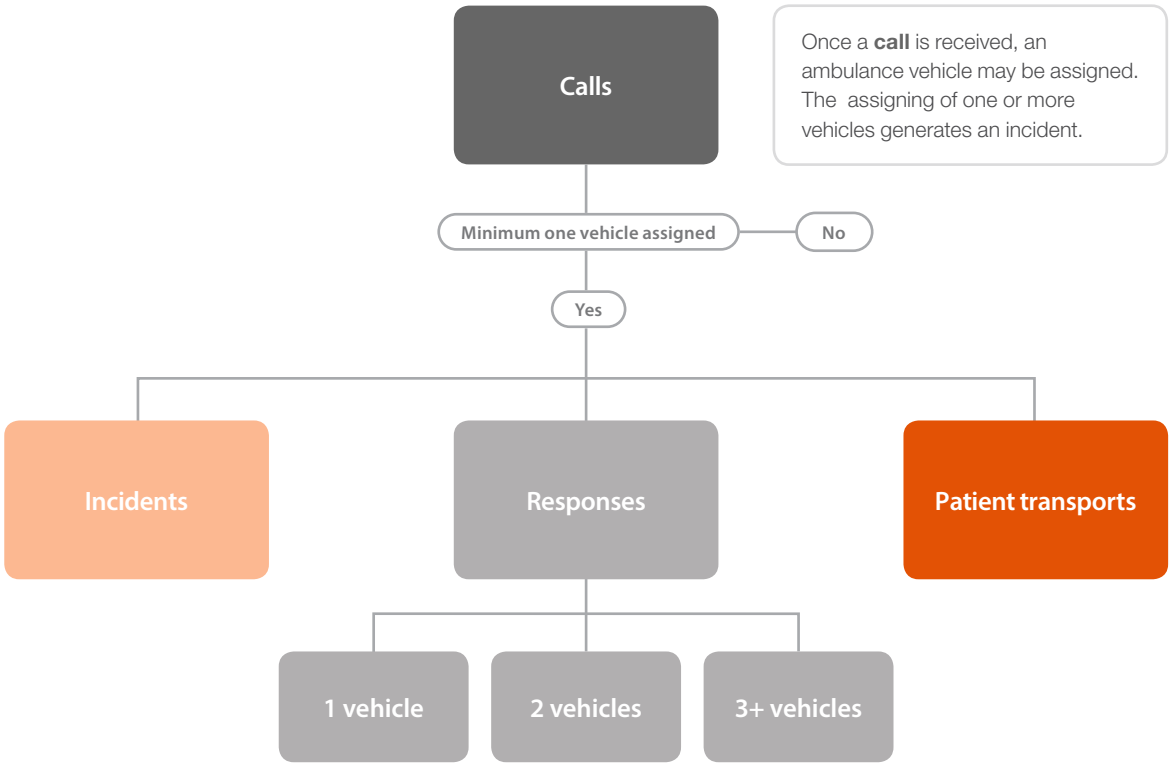
Units of measurement

Various units of activity can be used as the basis for ambulance performance assessment, including calls, responses, incidents or transports (Figure 12).

In NSW, a call defines the initiation of a cycle and is the point at which a dispatcher assesses the need for an ambulance. An incident occurs when an ambulance is assigned and dispatched to a scene. Multiple responses can result from a single incident if more than one ambulance is dispatched. Ambulance transports occur if a patient requires further medical care at a hospital.

Some jurisdictions base certain metrics on calls, particularly focusing on the time it takes to answer. In the NSW context, this is an operational measure used routinely by NSW Ambulance.

Figure 12 Units of measurement in ambulance activity, NSW



An **incident** is an event that results in a response by one or more ambulances.

Not all calls result in an ambulance response. Incidents are prioritised based on the urgency of the case.

A **response** is the dispatch of an ambulance from a local response area. Depending on the seriousness of an incident, or the number of people involved, multiple responses (vehicles) may be required for a single incident. Most incidents have one vehicle assigned. Around one in five has multiple vehicles assigned. Some vehicles are cancelled en route.

Some incidents involve one **patient** and some involve multiple patients.

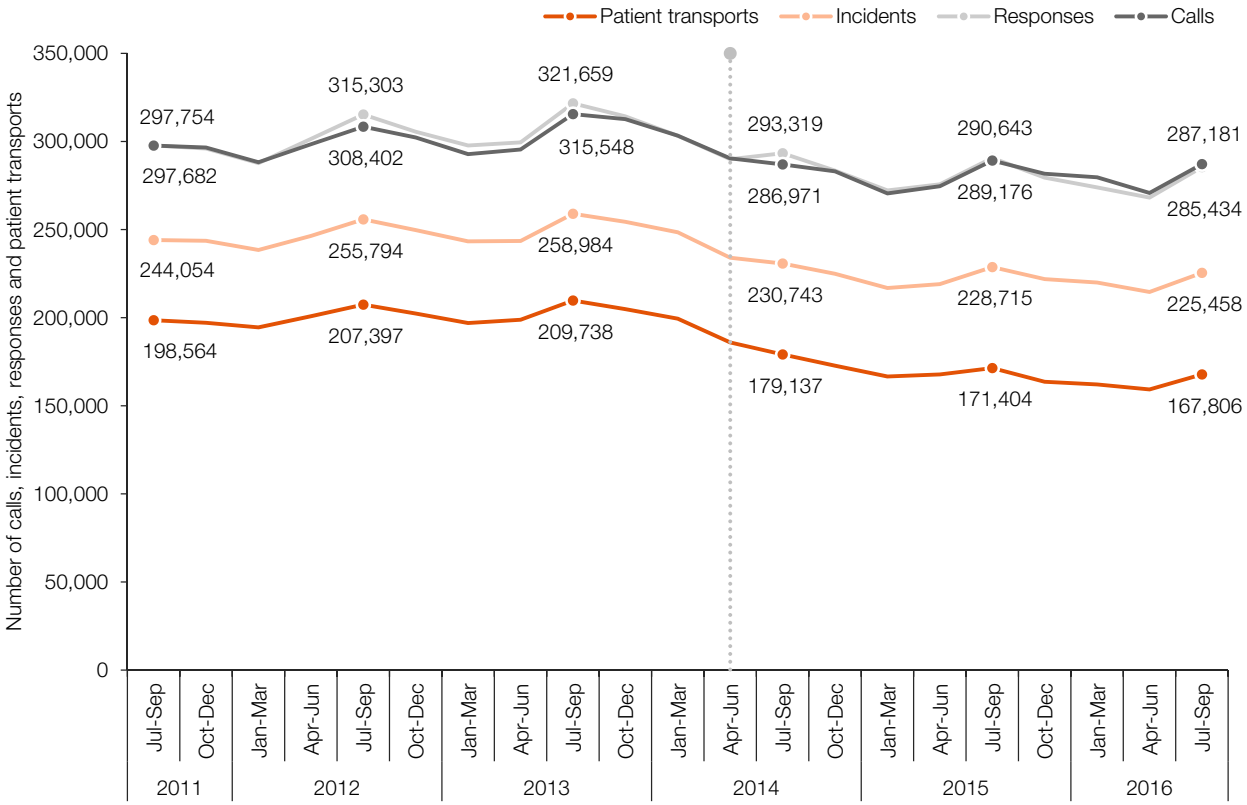
Once the ambulance has arrived at the scene, patients are treated and transported, or treated at the scene only. Most patients are **transported**.

Responses are the most commonly established base unit for timeliness measurement both nationally and internationally. They are highly relevant, important in operational terms and reflect patient demand.

Responses is the unit of measurement that will underpin ambulance performance reporting in *Healthcare Quarterly*.

Since April to June 2014, the number of calls, responses, incidents and patient transports have all decreased – largely as a result of policy changes, the introduction of telephone advice lines and alternative arrangements for transporting non-critical patients (such as the Non-Emergency Patient Transport service which was introduced in June 2014) (Figure 13).

Figure 13 Trends in measures of ambulance activity, NSW, July 2011 to September 2016



Capturing events of interest

Relevance of measures

Electronic information systems collect a range of data points that mark particular events or time intervals in an ambulance cycle.

There are 10 distinct time stamps recorded in the ambulance data. In many jurisdictions, metrics are often based on composite time periods that span a number of time points in the cycle (Figure 14).

Eight different permutations or 'events of interest' were explored in terms of relevance and face validity (Figure 15).

Seven time intervals were analysed for their sensitivity to variation in performance at two organisational levels (zones, and LRAs). Four of these – call to ambulance arrival time, mobilisation time, response time and turnaround time – were subject to further assessment of variation over time (pages 25-32).

Figure 14 Time stamps of the ambulance cycle, NSW

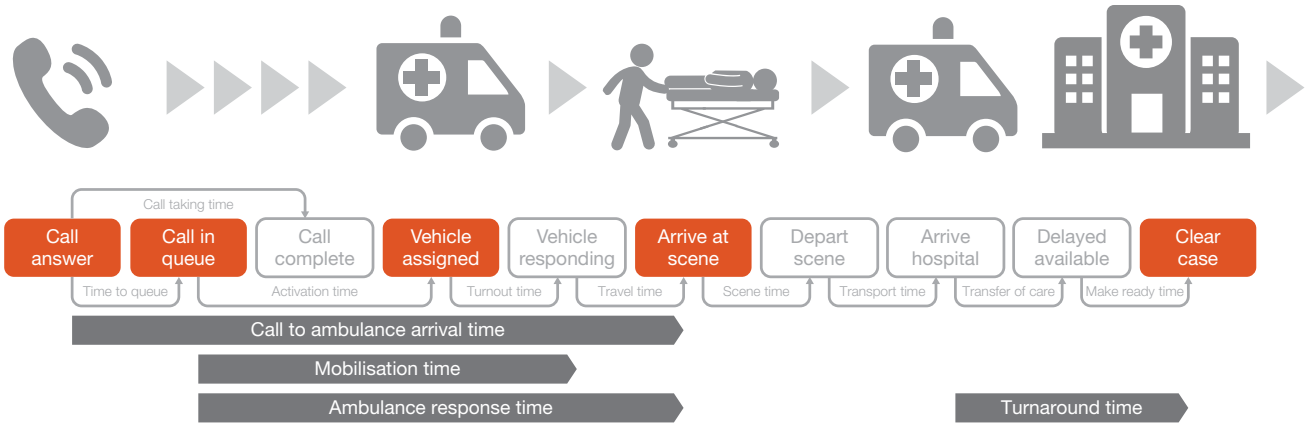


Figure 15 Performance indicators considered for public reporting in NSW

Measure	Description	Insights
Call to queue	The time from when a call is answered at the ambulance control centre to the time the call is placed in the 'electronic queue' for an ambulance to be dispatched.	A measure of telecommunications and the electronic dispatch system, this is an important operational measure. It deals with very small time intervals and can be considered as part of a longer time interval. In a NSW context, it does not require separate public performance reports.
Response time	From the time an emergency services operator takes the patient's details, determines a response is needed and places a call in the 'electronic queue' for an ambulance to be dispatched, to the time the first ambulance arrives on the scene.	This is a widely-used metric for ambulance performance reporting, however definitions differ slightly across jurisdictions (e.g. some start with patient call, others with call in queue). The time it takes for the ambulance service to arrive on the scene following an emergency call for help is a highly salient measure with considerable face validity. Response time can be affected by non-modifiable factors such as topography, distance and traffic conditions and these should be considered in performance assessment and attribution.
Mobilisation time	From the time an emergency call is placed in queue for an ambulance to the time the first vehicle is en route to the incident.	A measure of operational preparedness, mobilisation time is affected by organisational and staffing arrangements. For example, 24-hour LRAs generally have paramedics onsite and can mobilise relatively quickly, while non-24-hour LRA's mobilisation times are affected by the time it takes for paramedics to reach the ambulance vehicle before going to the scene of an incident.
Travel time	The time an ambulance takes to drive to the scene of an incident.	Travel time is shaped by non-modifiable factors such as distance, traffic conditions and topography. Alone, it is not reflective of performance yet it is an important component to overall responsiveness. It can be a meaningful measure of performance at a system level when reflecting on operational matters such as LRA location.
Turnaround time	From the time an ambulance which is transporting a patient arrives at a hospital to when the ambulance is ready to respond to a new incident.	Turnaround time captures the time an ambulance spends at a hospital. It is affected by inter-organisational transfer of patient care from ambulance to the emergency department staff as well as the time it takes for the vehicle to be cleaned and prepared for the next incident.
Make ready time	From the time a crew returns to the ambulance after transferring patient care to the emergency department, to when the ambulance is ready to respond to a new incident.	Make ready time is a subset of turnaround time and is not consistently recorded by ambulance crews. In some LRAs, crews 'double click' to record 'delayed available' and 'clear case' status almost simultaneously.
Call to ambulance arrival time	From the time a call is first answered in the ambulance control centre, to the time the first responding ambulance resource arrives at the scene of an incident.	Designed to capture timeliness from a patient perspective, call to ambulance arrival time reflects the time from when a call for assistance is answered to the time the first paramedics arrive on the scene. It is affected by travel time.
Case cycle time*	From the time an ambulance is assigned to an incident, to the time the case is cleared and the vehicle is ready for its next assignment.	Comprised of multiple time intervals that together form the entire ambulance journey. While case cycle time is useful for operational and planning purposes, it is difficult to clearly attribute performance to units. For performance assessment purposes, there is greater face validity and actionability when case cycle time is considered in specific sub-intervals.

* There is no information on case cycle time featured on the following pages. NSW Ambulance reports percentage of responses with priority 1 case cycle time within 60 minutes. In April 2016 this was 15.7% for NSW and ranged across sectors from 6.2% in Western Sydney-Nepean Blue Mountains to 42.7% in Western Sydney. Definitions for case cycle times differ for non-emergency cases and the clock starts with 'vehicle assigned'.

Capturing events of interest

Exploring options for timeliness measures

An essential characteristic of a performance measure is its ability to discriminate between meaningful and spurious variation. Seven different time intervals were assessed for discriminatory power. Three of the seven indicators, while important parts of the ambulance journey, were not found to be meaningful standalone measures of performance.

These three candidate measures had significant limitations in terms of minimal observed variation (call to queue), the limited opportunities for ambulance crews to improve performance (travel time), or imprecise measurement (make ready time) (Figures 16, 17 and 18). These intervals do however represent an important component of ambulance responses and are included in the composite time intervals of call to ambulance arrival time, response time and turnaround time.

The intervals with meaningful variation at the zone and LRA level were mobilisation time, turnaround time, response time and call to ambulance arrival time.

Priority category appears to influence the extent of variation for mobilisation time, response time and call to ambulance arrival time. For all three measures, median times get progressively longer and the range between lowest and highest median times widens as priority or urgency decreases (Figures 17 and 18).

For further analyses of variation in performance at zone and LRA levels over time see pages 25–32.

Figure 16 Zone-level variation in median time interval measures, NSW, July to September 2016

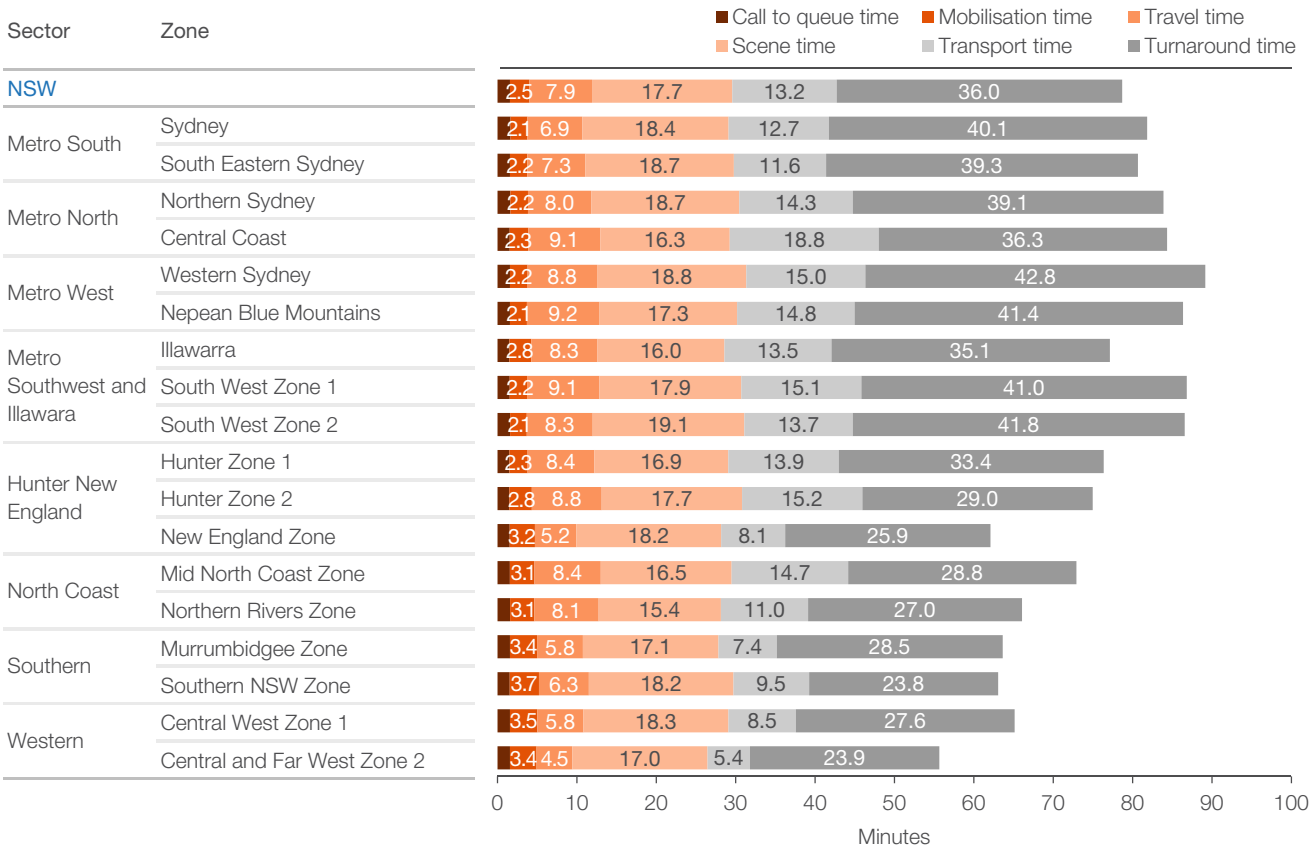


Figure 17 Exploring zone-level variation in time interval measures, NSW, July to September 2016

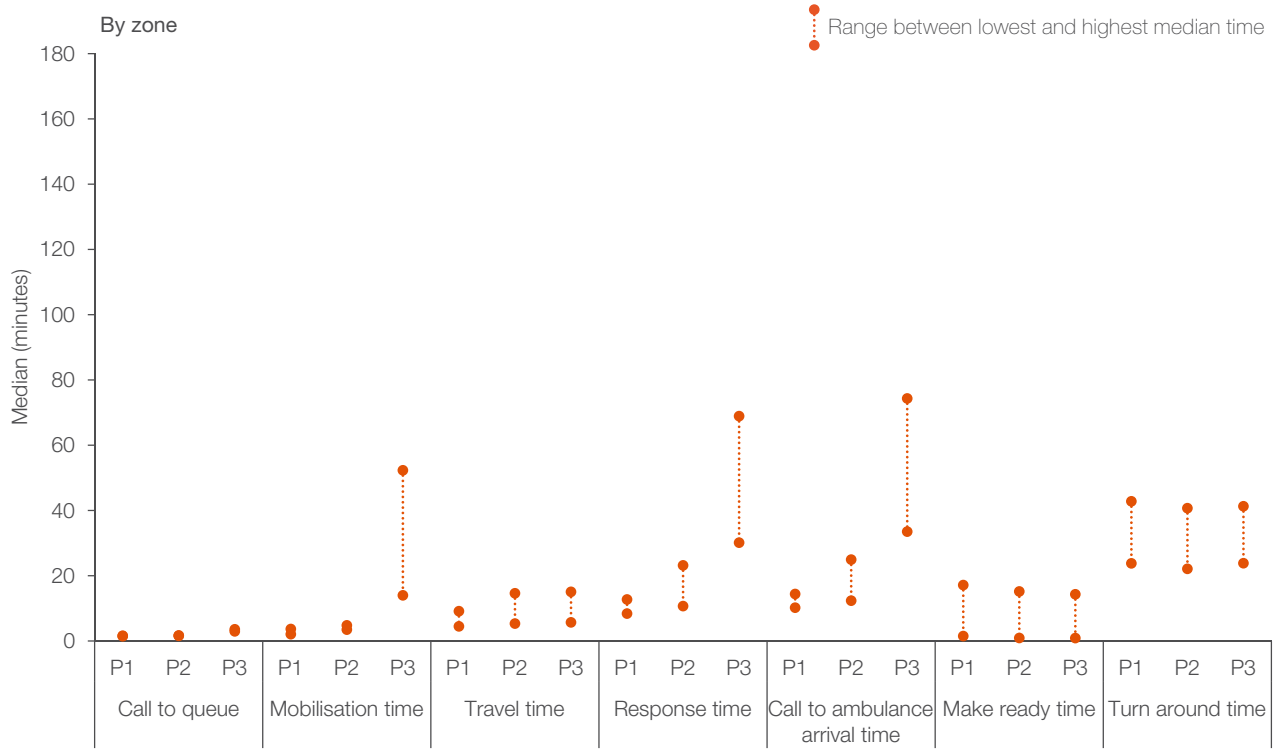
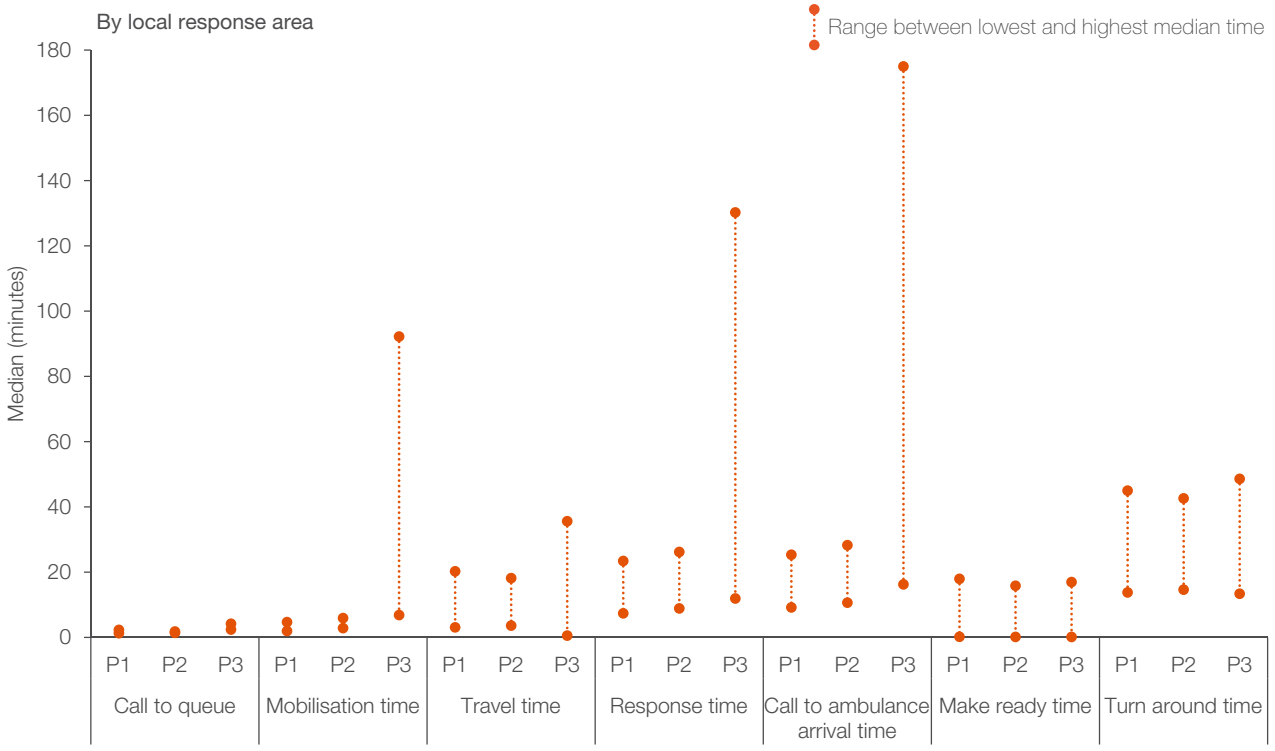


Figure 18 Exploring local response area-level variation in time interval measures, NSW, July to September 2016



Capturing events of interest

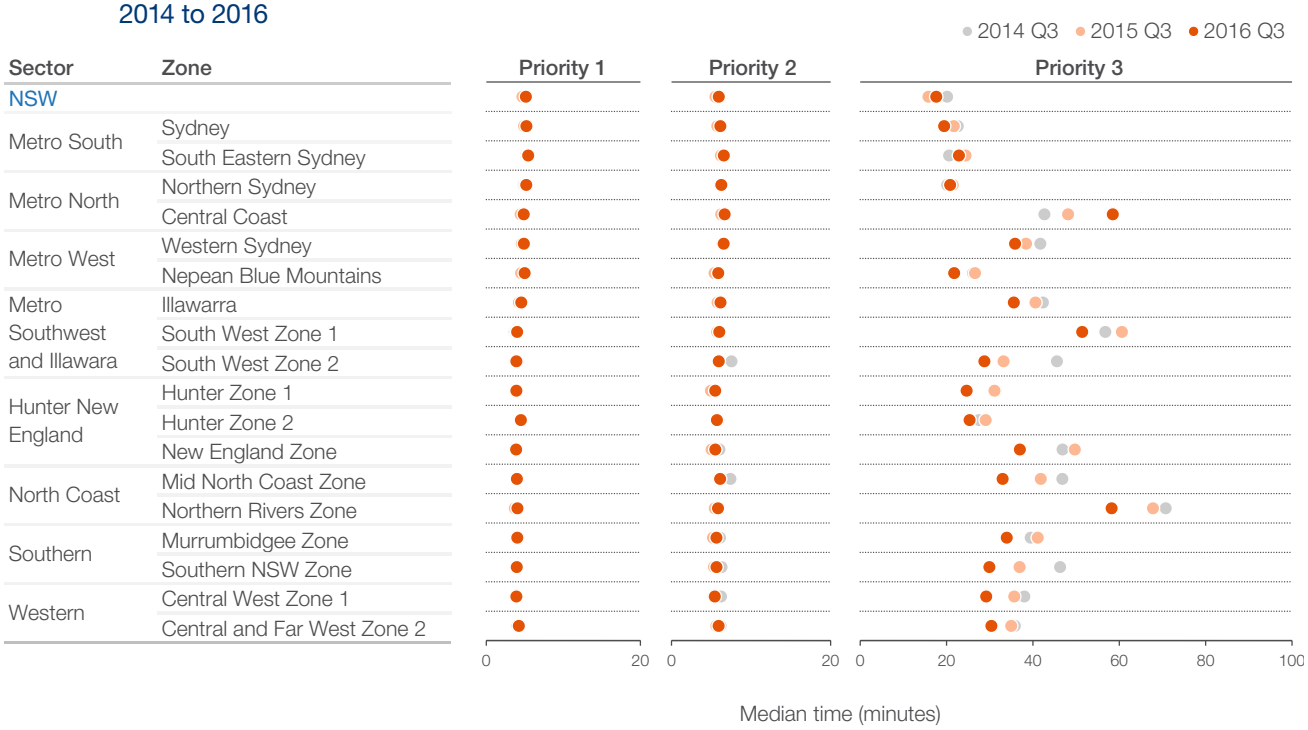
Assessment of mobilisation time

Once an ambulance has been given the details of a new call, there is typically a short period of time before the vehicle is en route, referred to as mobilisation time (Figure 19). Mobilisation time is a measure of preparedness and operational responsiveness.

This indicator can be affected by staffing and organisational arrangements. For example, responses from LRAs with on call or non-24-hour services often have longer mobilisation times as staff must come from outside the ambulance station to respond. As a result, median mobilisation time tends to be longer in LRAs in non-metropolitan zones compared to metropolitan zones [data not shown].

Figure 19 Intervals covering mobilisation time, NSW

Figure 20 Zone-level variation in median mobilisation time, by priority code, NSW, July to September quarters 2014 to 2016

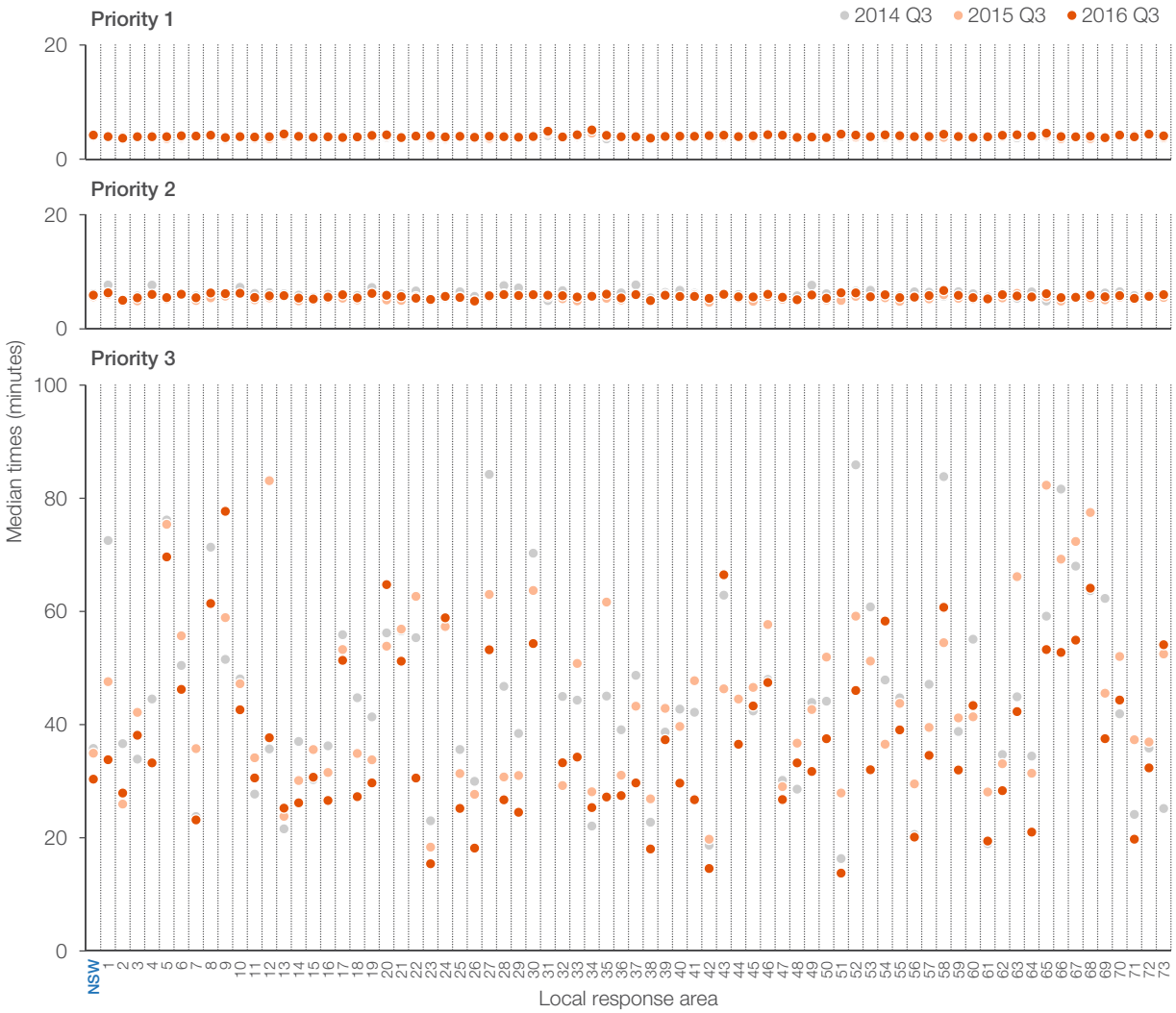


There are small differences between zones in quarterly results for priority 1 and 2 responses. Variation increases between zones, and over time for priority 3 (Figure 20).

For LRA results, there is little variation between units and over time for priority 1 and 2. In contrast, results for priority 3 show considerable variation reflecting time heterogeneity of priority 3 responses (Figure 21).

These analyses support the publication of mobilisation times (median and 90th percentile) in *Healthcare Quarterly* for priority 1 and 2 only.

Figure 21 Local response area-level variation in median mobilisation time for 24-hour local response area type, by priority code, NSW, July to September quarters 2014 to 2016



Capturing events of interest

Assessment of turnaround time

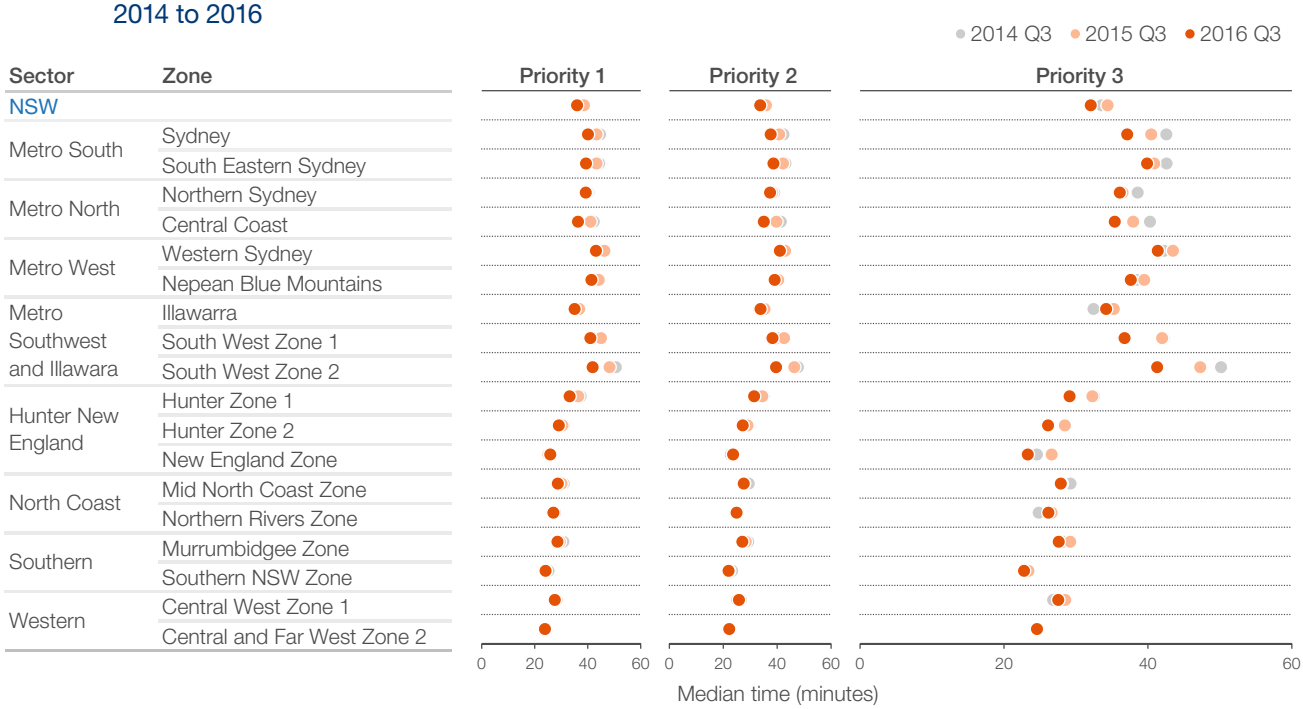
Turnaround time refers to the period paramedics spend at hospital emergency departments. It is measured from the time an ambulance transporting a patient arrives at a hospital to when it is 'clear' and ready to respond to another incident (Figure 22).

The pattern of median times across zones for priority 1 and 2 has a similar profile. For priority 3, there is greater variation in median turnaround time (Figure 23).

At the zone level, there is almost twofold variation in median priority 1 turnaround times – ranging in 2016 from 21 to 40 minutes (Figure 23).

Figure 22 Intervals covering turnaround time, NSW

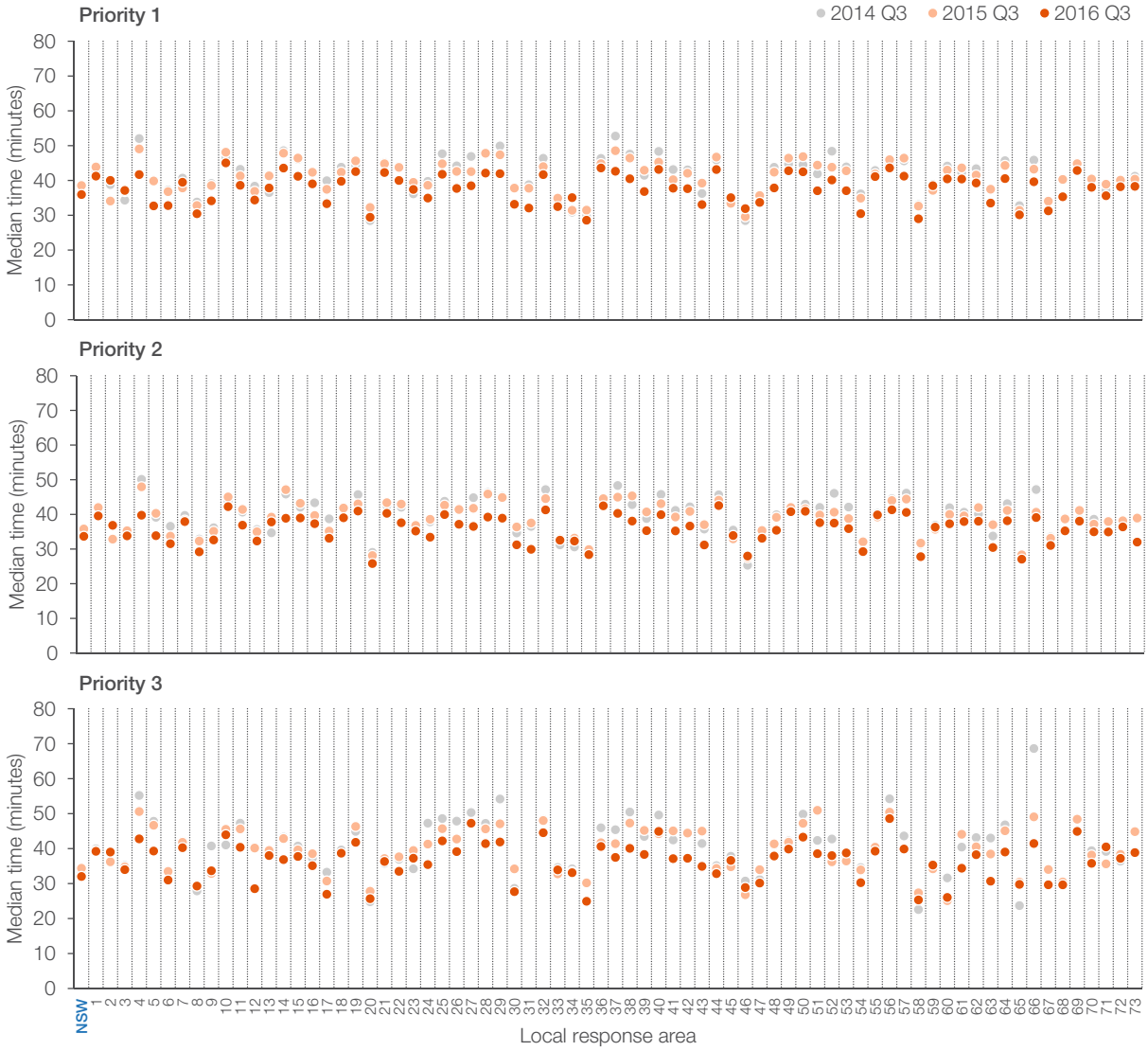
Figure 23 Zone-level variation in median turnaround time, by priority code, NSW, July to September quarters 2014 to 2016



Data from 24-hour LRAs show similar patterns of variation across the priority categories, albeit with amplified variation for priority 3 (Figure 24).

These analyses support the publication of turnaround times (median and 90th percentile) in *Healthcare Quarterly*.

Figure 24 Local response area-level variation in median turnaround times for 24-hour local response areas, by priority code, NSW, July to September quarters 2014 to 2016



Capturing events of interest

Assessment of response time

The time it takes for an ambulance to arrive on the scene following an emergency call for help is a highly salient measure with considerable face validity. As a result, response time is probably the most widely used metric for ambulance performance reporting internationally.

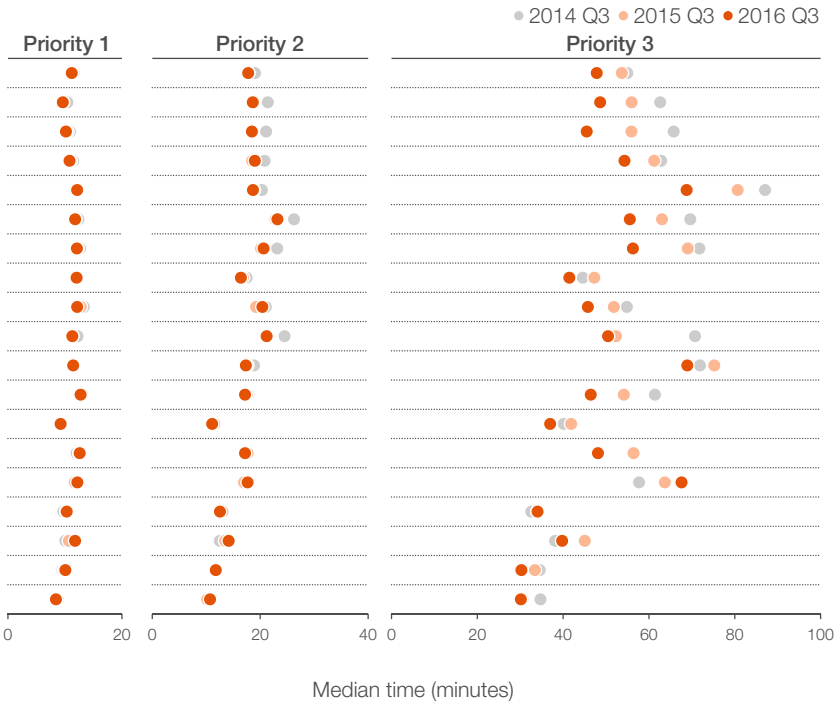
Definitions differ slightly across jurisdictions, however in NSW the clock starts when the call is placed in the queue for an ambulance operator and stops with arrival at the scene (Figure 25).

There are only small differences in quarterly results at a zone level for priority 1 responses. Variation increases slightly both between zones and over time for priority 2 and markedly so for priority 3 (Figure 26).

Figure 25 Intervals covering response time, NSW

Figure 26 Zone-level variation in median response time, by priority code, NSW, July to September quarters 2014 to 2016

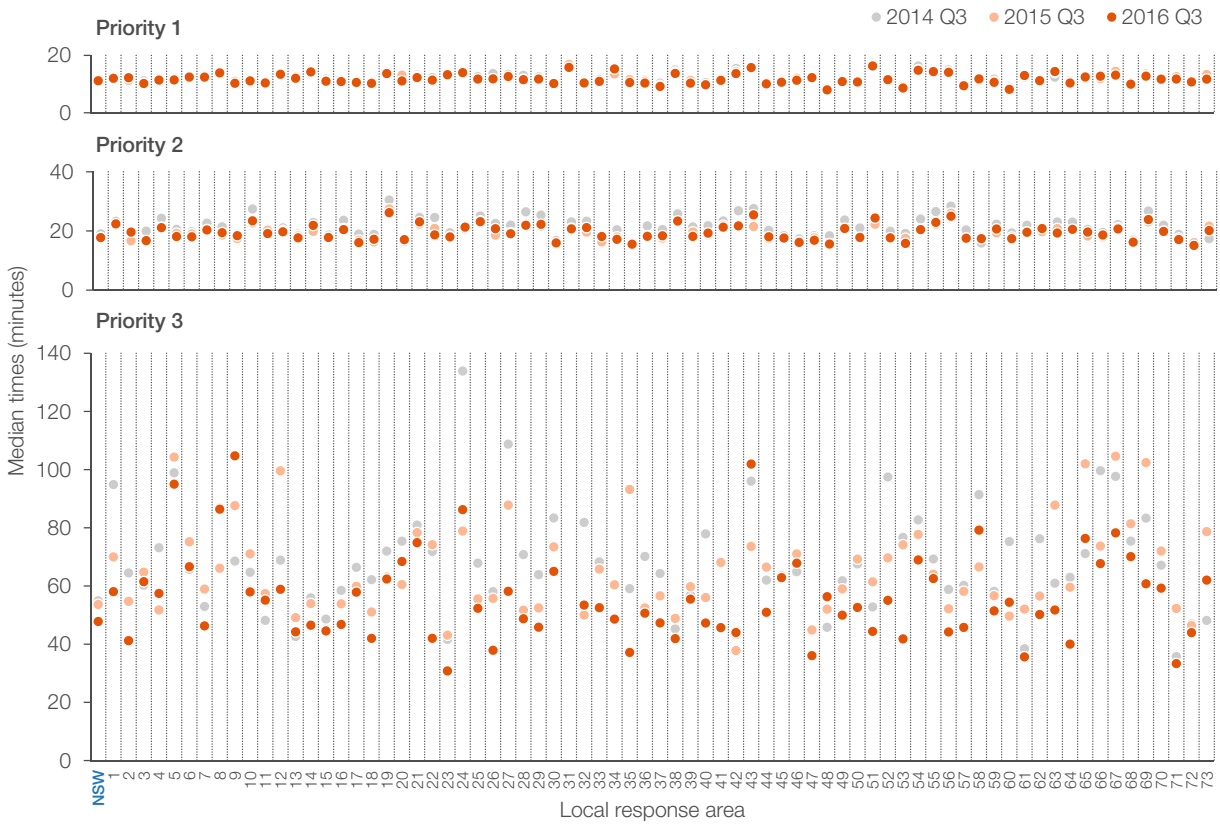
Sector	Zone
NSW	
Metro South	Sydney
	South Eastern Sydney
Metro North	Northern Sydney
	Central Coast
Metro West	Western Sydney
	Nepean Blue Mountains
Metro Southwest and Illawara	Illawarra
	South West Zone 1
	South West Zone 2
Hunter New England	Hunter Zone 1
	Hunter Zone 2
	New England Zone
North Coast	Mid North Coast Zone
	Northern Rivers Zone
Southern	Murrumbidgee Zone
	Southern NSW Zone
Western	Central West Zone 1
	Central and Far West Zone 2



At the LRA level, data from 24-hour type stations shows there is variation both between units and over time in all priority categories. The results for priority 3 show considerable heterogeneity (Figure 27).

These analyses support the publication of response times for priority 1 and priority 2 (median and 90th percentile) in *Healthcare Quarterly*.

Figure 27 Local response area variation in median response times for 24-hour local response area type, by priority code, NSW, July to September quarters 2014 to 2016



Capturing events of interest

Assessment of call to ambulance arrival time

Call to ambulance arrival time is a metric that aims to capture timeliness from a patient perspective. It reflects the time from when a triple zero call for assistance is answered to the time the first paramedics arrive at the scene (Figure 28).

There are only small differences in quarterly results between zones for priority categories 1 and 2 responses. Variation increases between zones and over time are more marked for priority 3 (Figure 29).

Figure 28 Intervals covering call to ambulance arrival time, NSW

Figure 29 Zone-level variation in call to ambulance arrival time, by priority code, NSW, July to September quarters 2014 to 2016

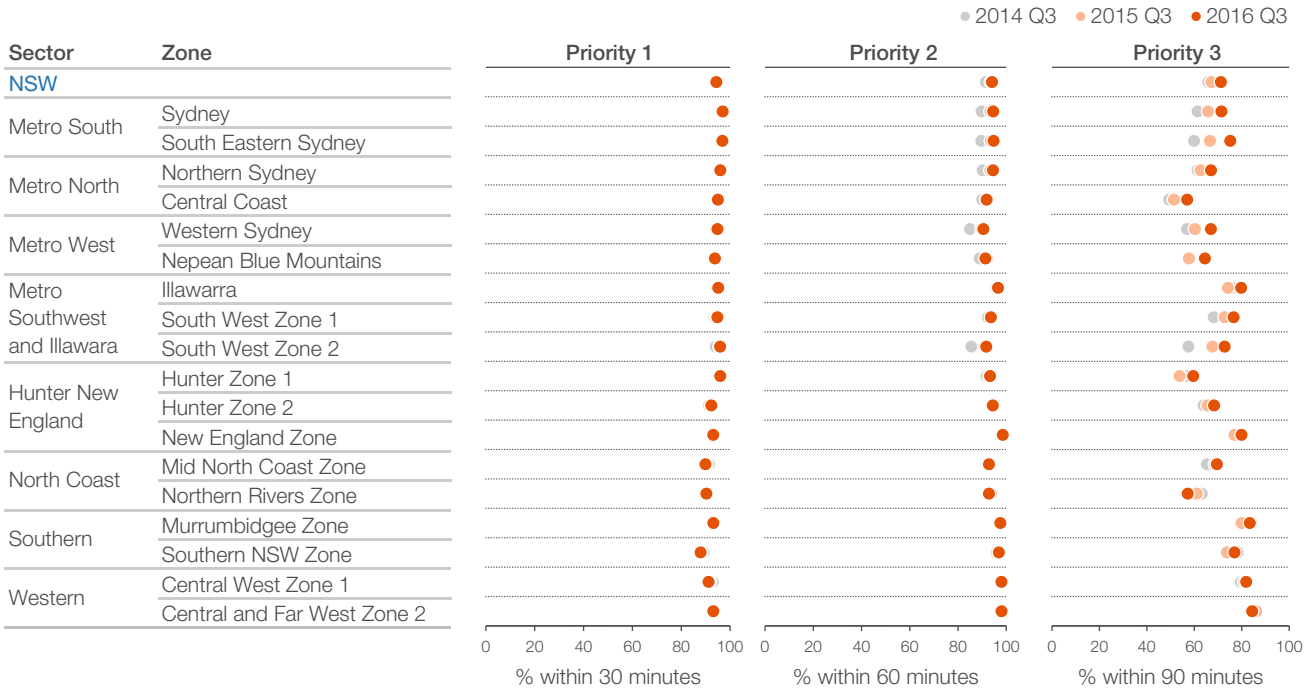


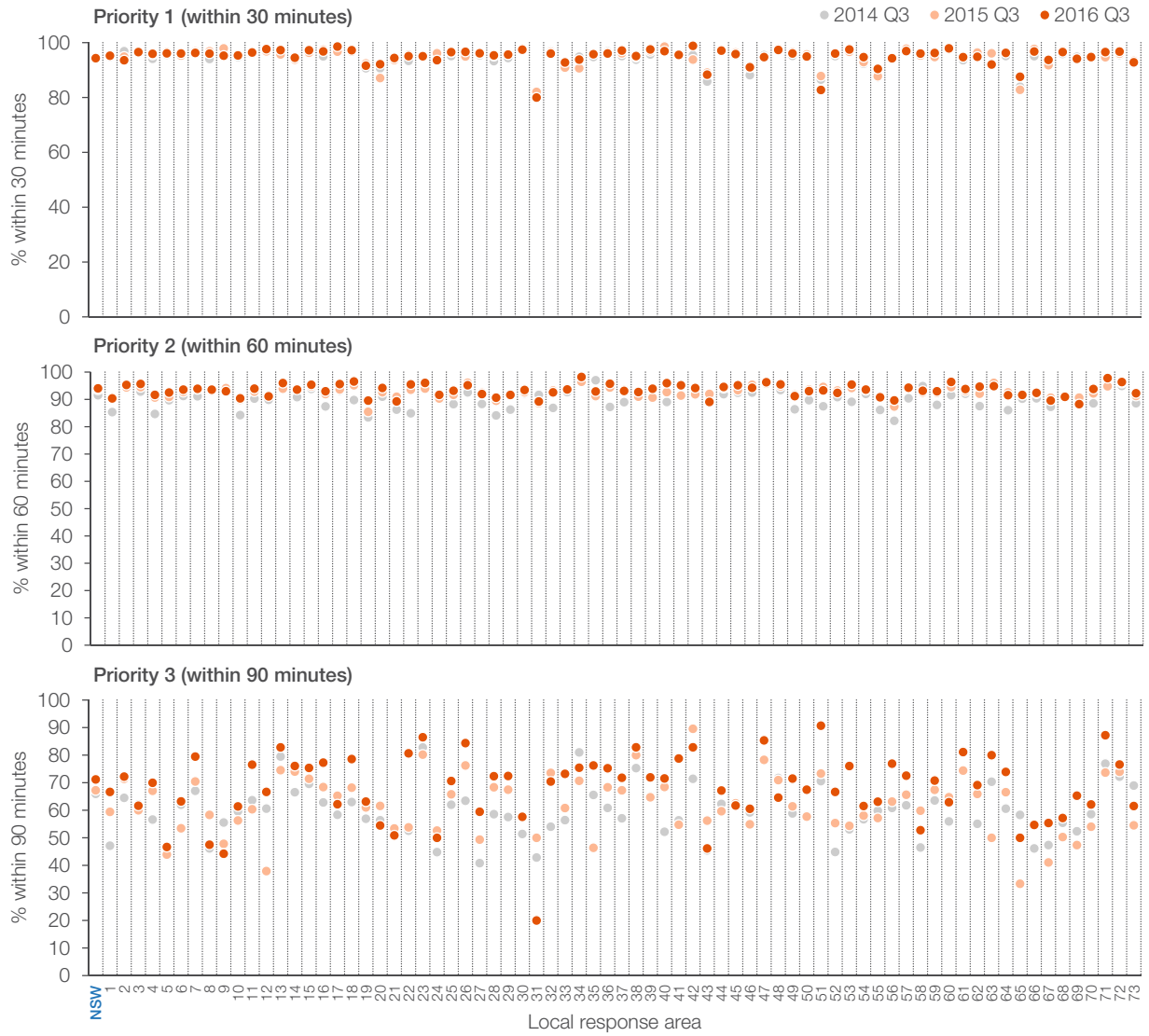
Figure 30 shows LRA-level results for call to ambulance arrival times (using the percentage completed within defined times).

Analysis of data from the 24-hour LRA type shows there is variation for priority 1 and priority 2 both between units and over time. The results for priority 3 show considerable heterogeneity (Figure 30).

These analyses support the publication of call to ambulance arrival times for priority categories 1 and 2, using the percentage within 30 minutes and 60 minutes, respectively. The heterogeneity of priority 3 and the unquantified impact of pre-booked responses precludes public reporting of priority 3 call to ambulance arrival times.

The setting of defined times and thresholds for call to ambulance arrival time metrics are further explored in pages 47–50.

Figure 30 Local response area variation in call to ambulance arrival times for 24-hour local response area type, by priority code, NSW, July to September quarters 2014 to 2016



Capturing events of interest

Comparing response time and call to ambulance arrival time

Response time and call to ambulance arrival time are closely related. With respect to time intervals, the measures differ only in terms of starting point. The clock for call to ambulance arrival time begins when the call for emergency services is answered, while the clock for response time starts when a request for an ambulance dispatch has been made ('call in queue'). For both measures, the clock stops when the first ambulance arrives at the scene.

Highly congruent in the intervals measured, when the same statistic is used (e.g. median), the two measures are strongly correlated (Figure 31). Empirically, this suggests that the information gained from each indicator is similar. However, using a normative standard to measure performance for call to ambulance arrival time (that is, the percentage of patients for whom an ambulance arrives within a defined time) differentiates the measures and provides a patient perspective – reflecting the actual time waited.

Figure 32 shows the relationship between median response times and percentage of call to ambulance arrivals within recommended times for the three priority categories. It shows that for priority 1 and 2 responses, median response times capture more variation. For priority 3 there is substantial variation for both median response time and the percent of call to ambulance arrivals within defined times.

Reporting both measures facilitates different approaches to benchmarking and time series measurement. Median response time is a well-established metric in the NSW context and is used by both the ambulance service and the NSW Ministry of Health for operational, planning and management purposes. Call to ambulance arrival time within 15 minutes for priority 1 calls is similar to measures reported in other jurisdictions, providing context interpret NSW results.

Figure 31 Correlation between median response time and median call to ambulance arrival time, priority 1 responses, NSW, July to September 2016

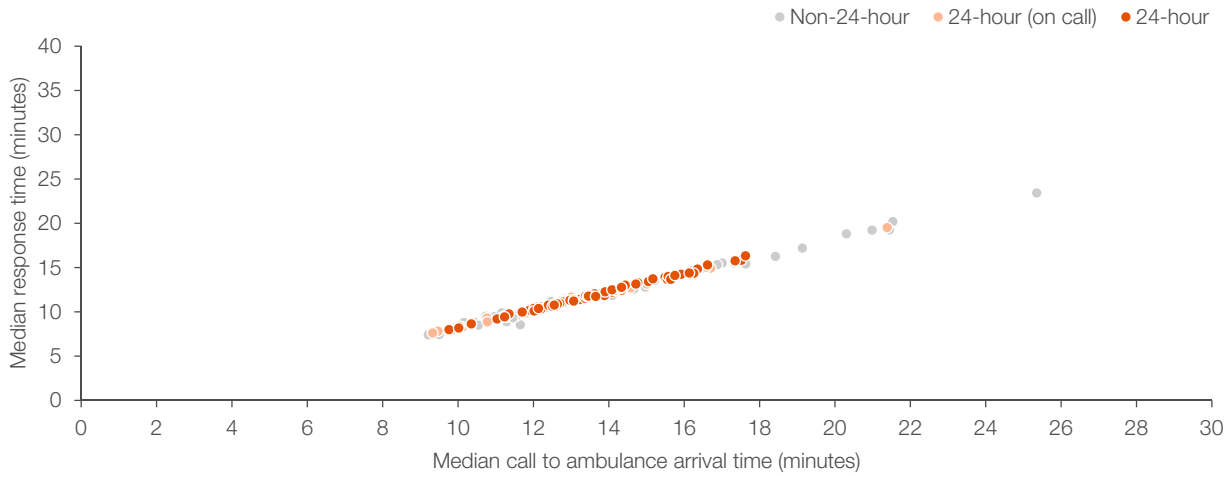
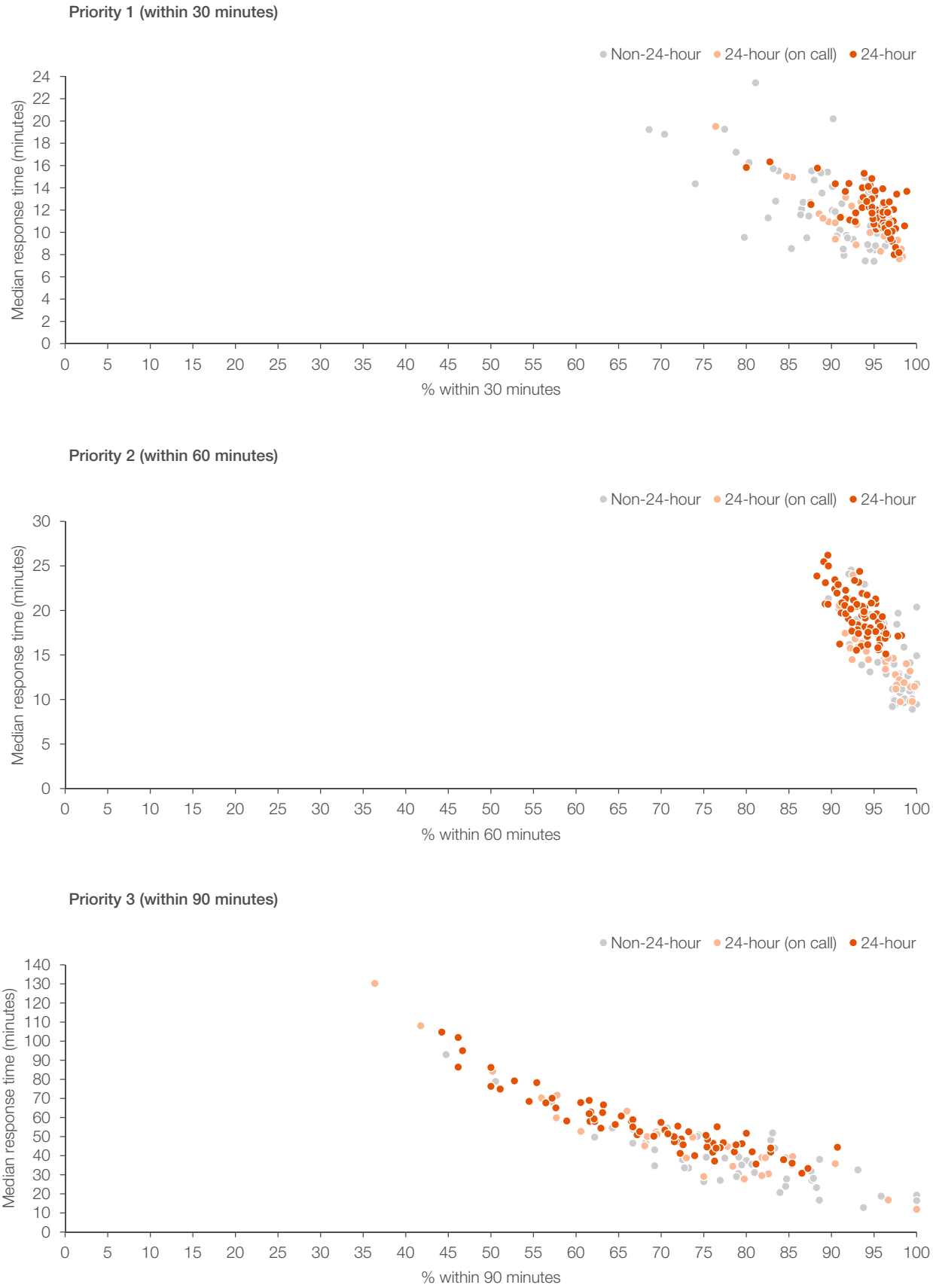


Figure 32 Correlation between median response time and percentage of responses within defined call to ambulance arrival time (priority categories 1, 2 and 3), NSW, July to September 2016



Capturing events of interest

Deconstructing call to ambulance arrival time

The call to ambulance arrival time comprises several time intervals and spans from the time a call is answered to the time the first ambulance arrives at the scene (Figure 33).

The median time for each interval making up call to arrival time was analysed and used to inform interpretation and indicator development (Figure 34 and 35).

Across time intervals (Figure 35), the least variation was seen in call to queue time (from the axis to the dark orange dot) while most variation is seen in travel time (between the light orange and the grey dots).

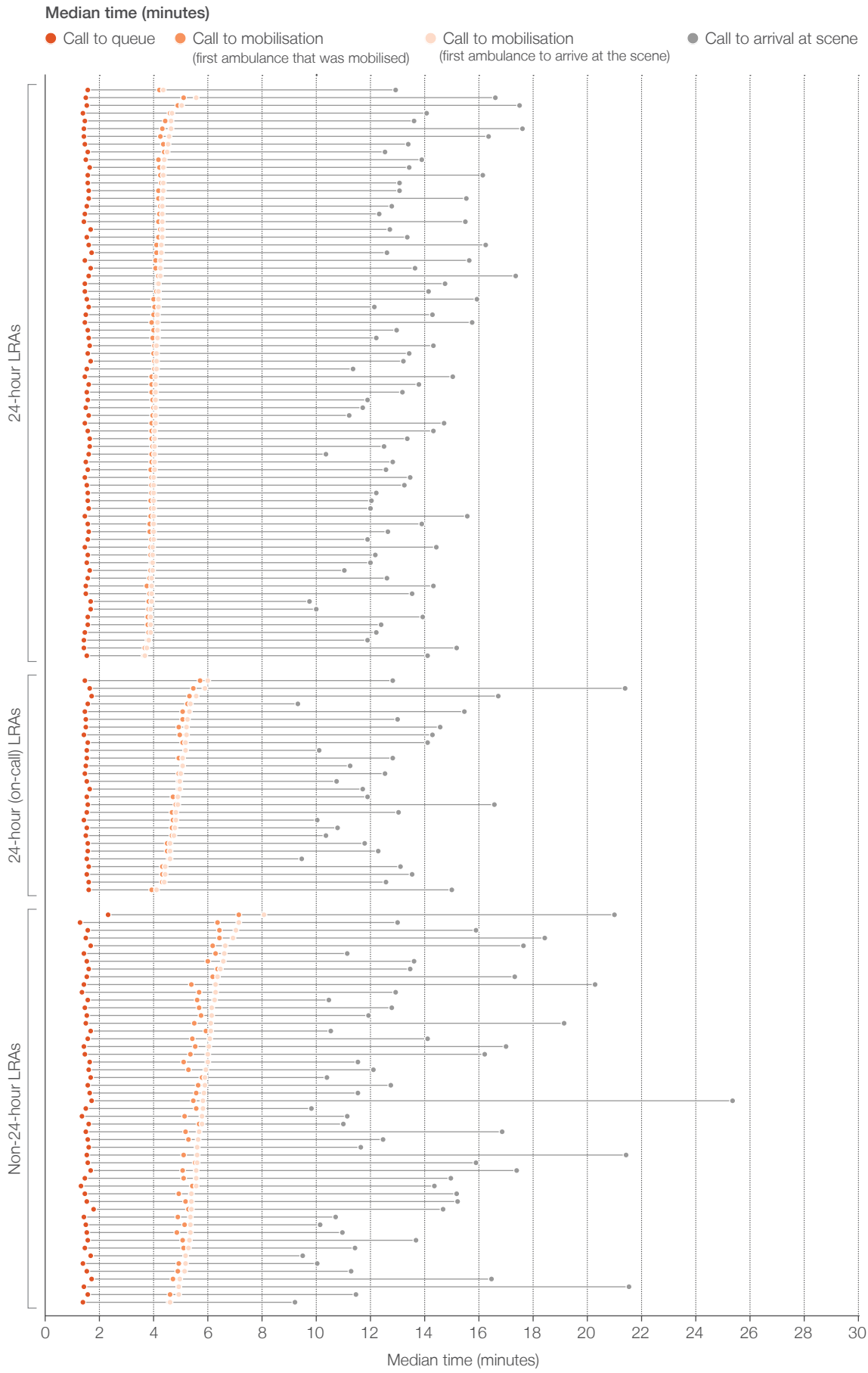
Figure 35 shows the relationship between the median mobilisation time (the orange dot) and the LRA type, highlighting the role that staffing and organisational arrangements have on mobilisation time. However, mobilisation time is not correlated with call to ambulance arrival time performance overall. LRA type does not appear to strongly predict call to ambulance arrival time.

Figure 33 Intervals covering call to ambulance arrival time, NSW

Figure 34 Median time (minutes) for the intervals in the call to ambulance arrival time, priority 1 responses, by local response area type, NSW, July to September 2016

24-hour	Min	Max	Range
Call to queue	1.4	1.7	0.3
Call to mobilisation (first ambulance that was mobilised)	3.7	5.1	1.4
Call to mobilisation (first ambulance to arrive at the scene)	3.7	5.6	1.9
Call to arrival at scene	9.8	17.6	7.8
24-hour (with on call)			
Call to queue	1.5	1.7	0.2
Call to mobilisation (first ambulance that was mobilised)	3.9	5.7	1.8
Call to mobilisation (first ambulance to arrive at the scene)	4.1	6	1.9
Call to arrival at scene	9.3	21.4	12.1
Non-24-hour			
Call to queue	1.3	2.3	1.0
Call to mobilisation (first ambulance that was mobilised)	4.6	8.1	3.5
Call to mobilisation (first ambulance to arrive at the scene)	4.6	7.2	2.6
Call to arrival at scene	9.2	25.4	16.2

Figure 35 **Median time for the intervals in the call to ambulance arrival time, priority 1 responses, by local response areas, and type, NSW, July to September 2016**



Attributing performance

Modifiable time intervals

Analyses explored the degree to which variation could be attributed to the local, zone or system level. Fundamental to this was distinguishing between modifiable and non-modifiable factors that affect performance.

Ambulance mobilisation time is a modifiable subset of call to ambulance arrival time that influences timeliness. NSW Ambulance measures the percentage of mobilisation times within three minutes for priority 1 responses.

There is a progressive upward shift in median mobilisation time on the metropolitan to non-metropolitan spectrum (Figure 33). Compared with other LRA types, 24-hour stations generally fall below a three minute median time. LRAs in the 24-hour (with on call) and the non-24-hour groups have median mobilisation times consistently above three minutes (Figure 36).

Analyses were conducted to assess whether a non-modifiable component, such as the median travel time, was correlated with median mobilisation time (i.e. whether zones with higher median mobilisation times in non-metropolitan areas also had higher median travel times).

Median travel time is more evenly distributed across geographic zones (Figure 38). Greater variation in travel time in more remote areas may be explained by the longer distances that ambulances often travel to reach incidents, and the extent to which ambulances respond to incidents in neighbouring LRAs.

Figure 36 Correlation of median mobilisation time and median travel time, local response areas, July to September 2016

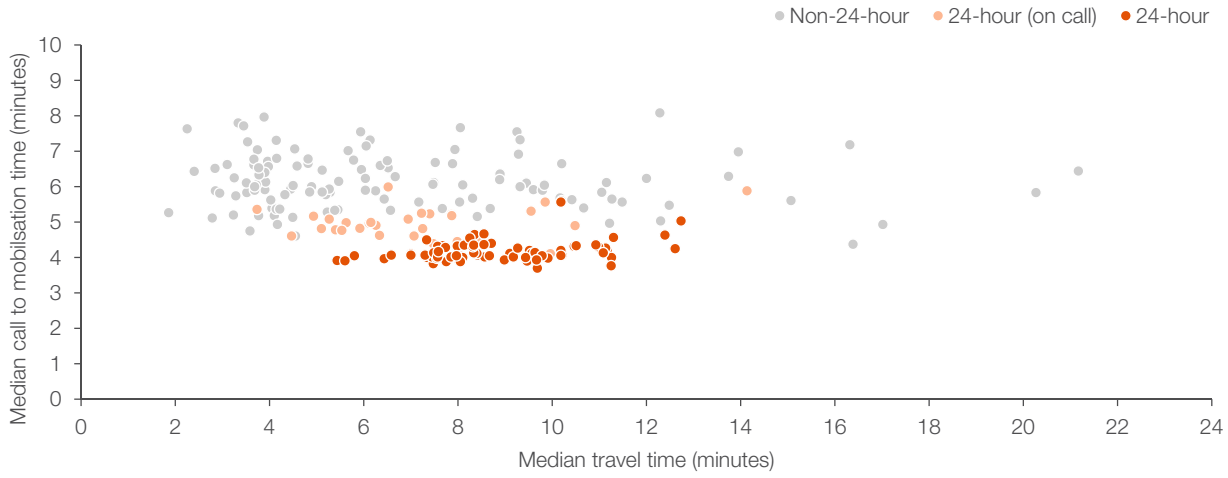


Figure 37 Distribution of priority 1 median mobilisation time by zone and local response area type, NSW, July to September 2016

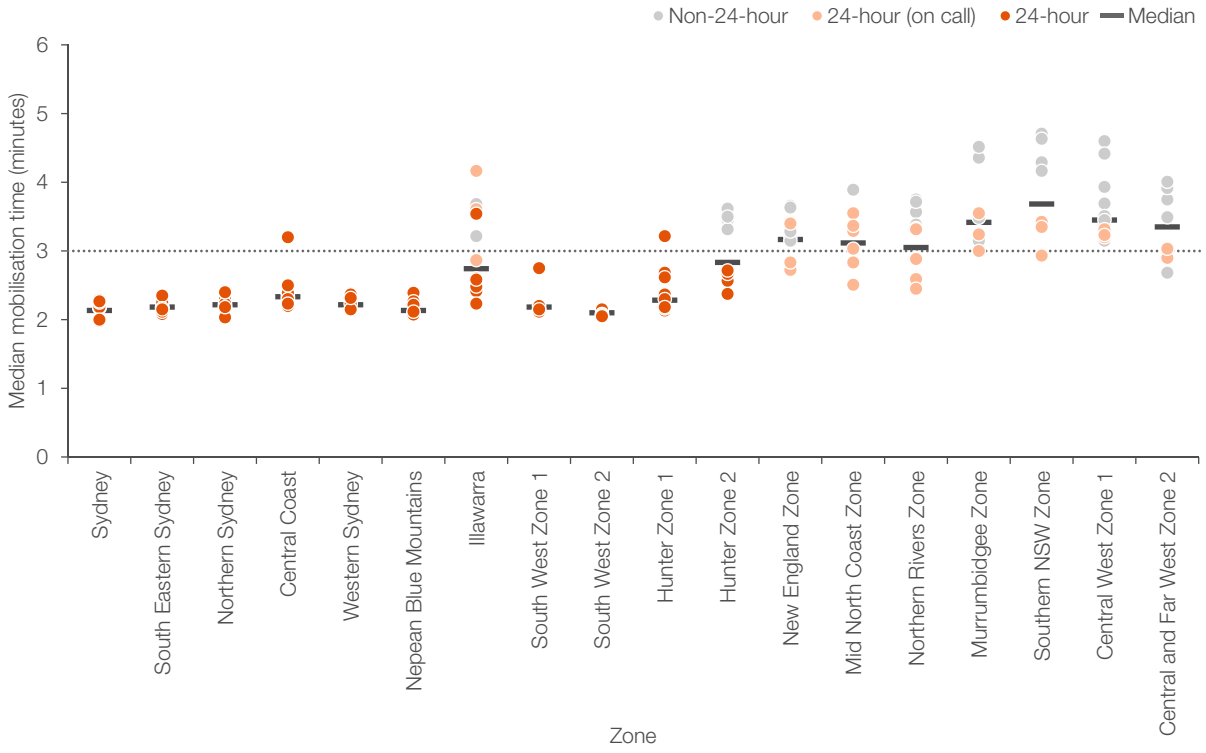
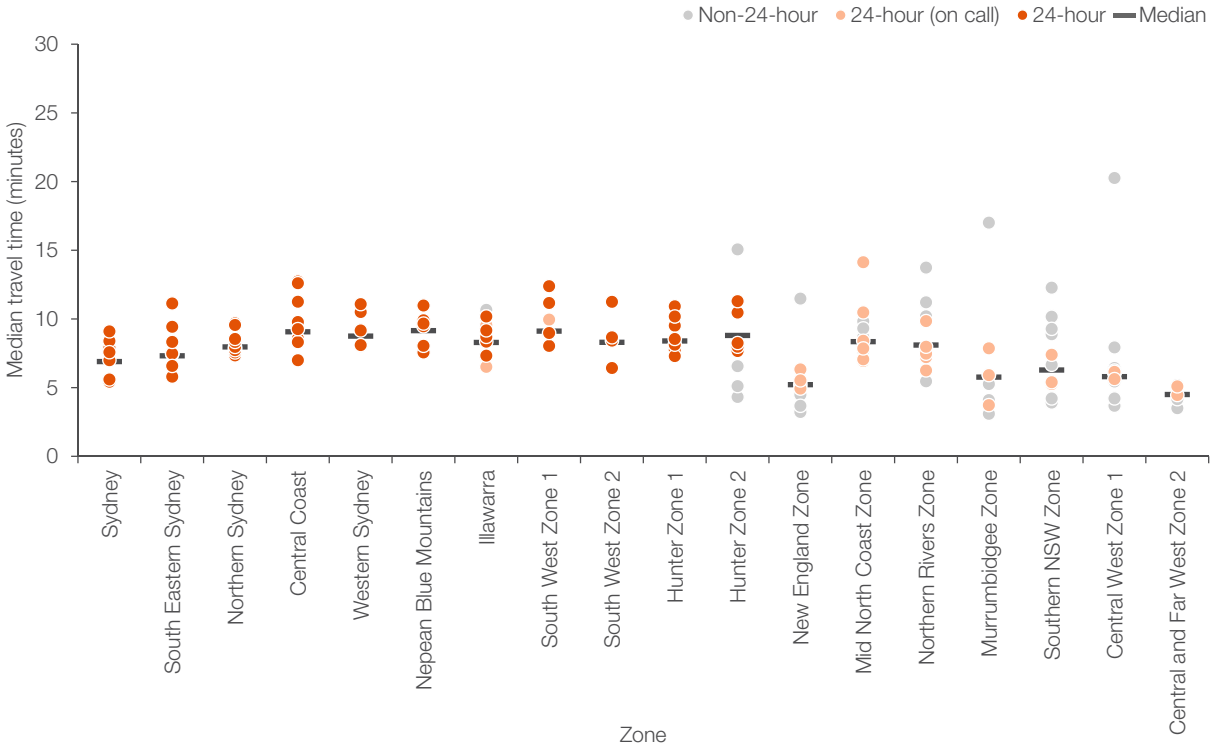


Figure 38 Distribution of priority 1 median travel times by zone and local response area type, NSW, July to September 2016



Risk adjustment

Making fair comparisons

In order to make fair comparisons, analyses are often stratified to take into account factors such as geographical, organisational or patient characteristics that are outside of the control of service providers.

In all LRA types, responses are predominantly either priority 1 or 2 (Figure 39).

The distribution of median response times is similar across the three different types of LRAs, suggesting that the type of LRA is not a meaningful basis for stratification of results (Figure 40).

Figure 39 Distribution of responses by local response area type and priority category, NSW, July to September 2016

Local response area	Responses
Peer group 1: 24-hour	189,730
Peer group 2: 24-hour (with on call)	51,292
Peer group 3: Non-24-hour	41,702
Peer group 4: Volunteer and community	2,013

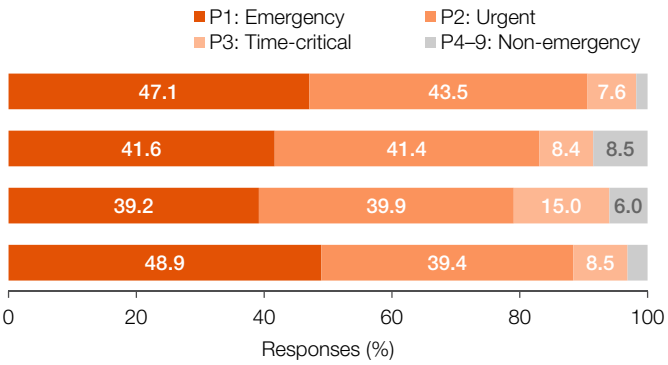
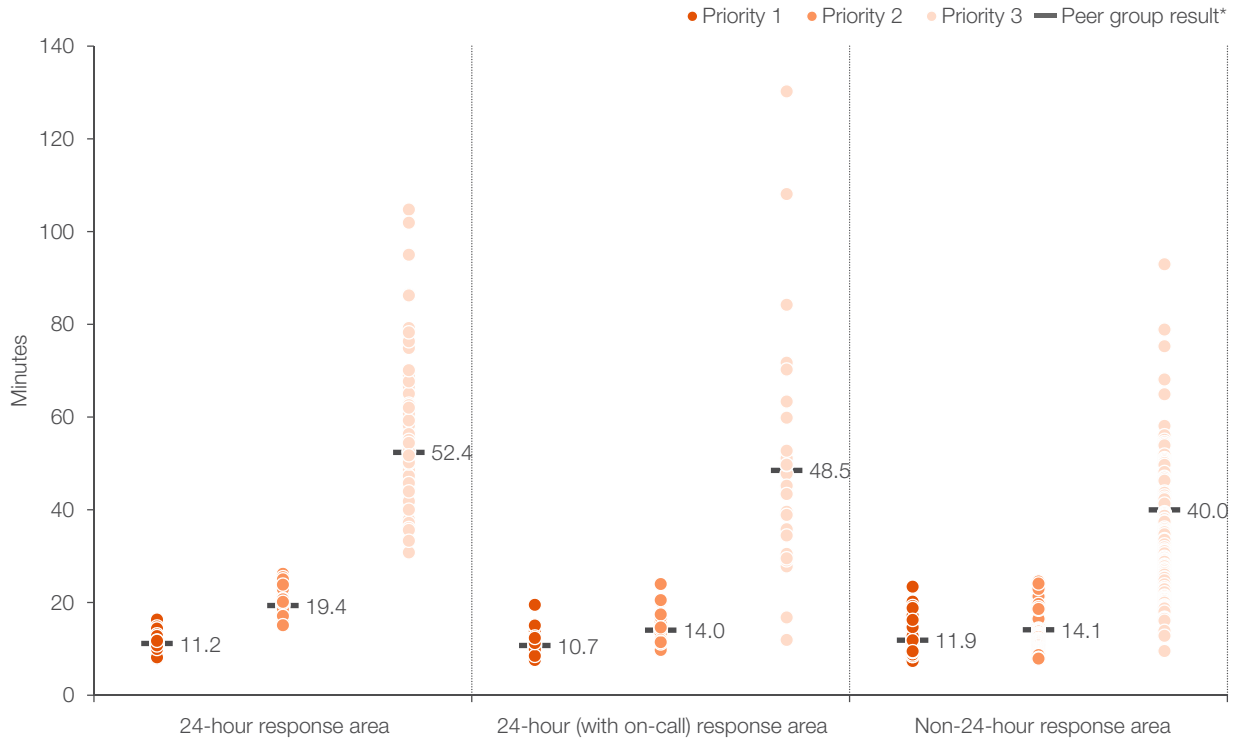


Figure 40 Median response time by local response area type and priority category, NSW, July to September 2016



24-hour on call and non-24-hour LRA types in non-metropolitan areas might be expected to have longer response times due to unmodifiable factors such as long travel times. If so, stratification by type could make allowances for the influence of travel time on performance. However, the distribution of response time that is spent on travel in each priority category appears similar across the LRA types (Figure 41).

Figure 41 Distribution of responses by priority category and local response area type, NSW, July to September 2016

Priority category	Responses
P1: Emergency	127,986
P2: Urgent	121,250
P3: Time-critical	25,220
P4–9: Non-emergency	10,281

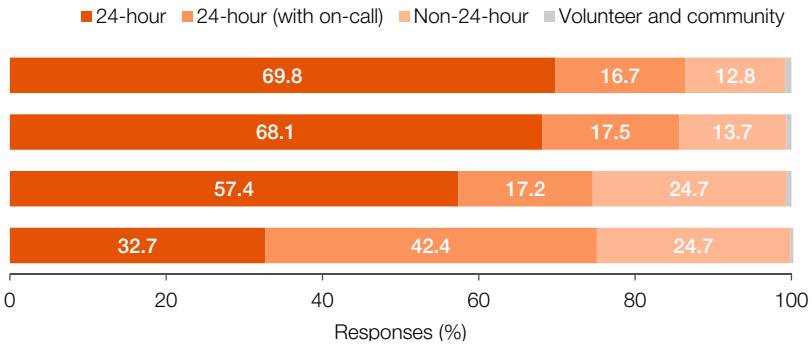
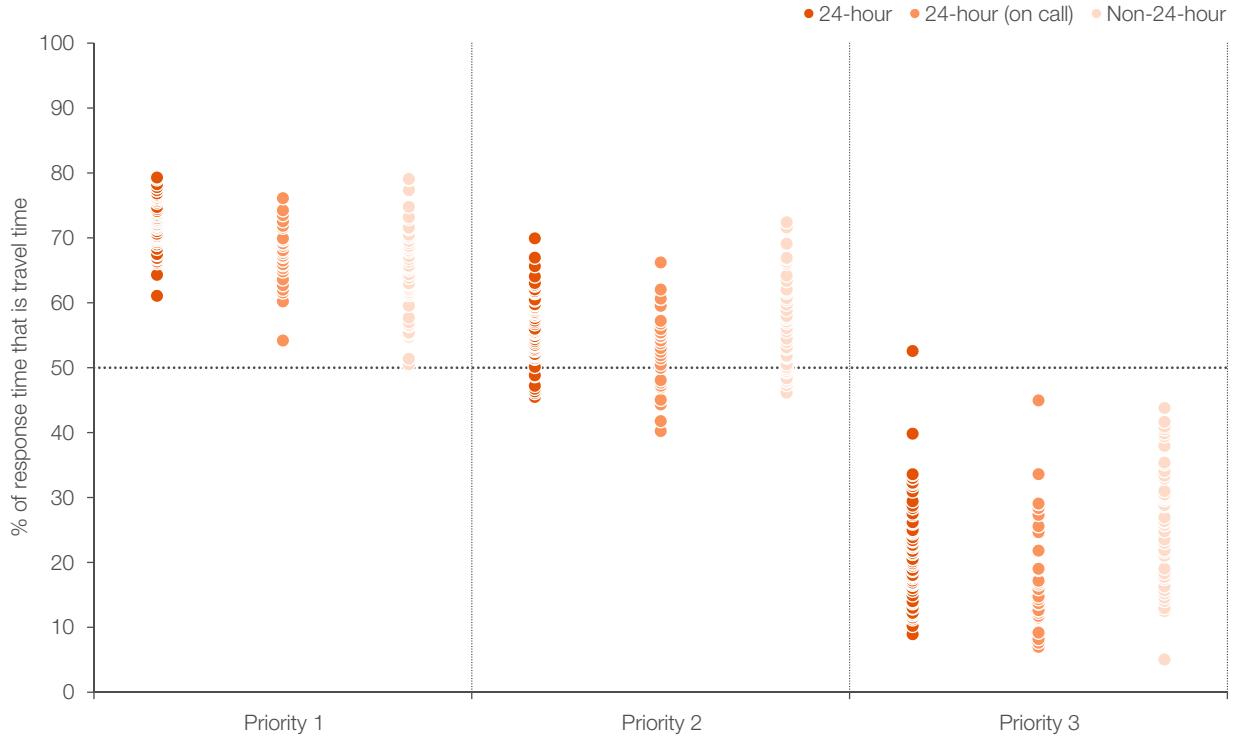


Figure 42 Percentage of response time that is travel time, priority 1 responses, by local response area type, NSW, July to September 2016



Risk adjustment

Variation in responses by priority

Assessments of variation in call to ambulance arrival time established that there is meaningful variation between zones, LRAs and over time for priority categories 1 and 2 (see pages 31-32).

An additional question is whether performance in each LRA is consistent across priority categories. This analysis shows for each LRA the percentage of call to ambulance arrival times within 30 minutes for priority 1; within 60 minutes for priority 2; and within 90 minutes for priority 3.

Results for the priority 1 and 2 responses are generally quite similar, but results for priority 3 responses vary considerably (Figures 43 to 45).

Figure 43 24-hour local response area call to ambulance arrival within defined times, by priority category, NSW, July to September 2016

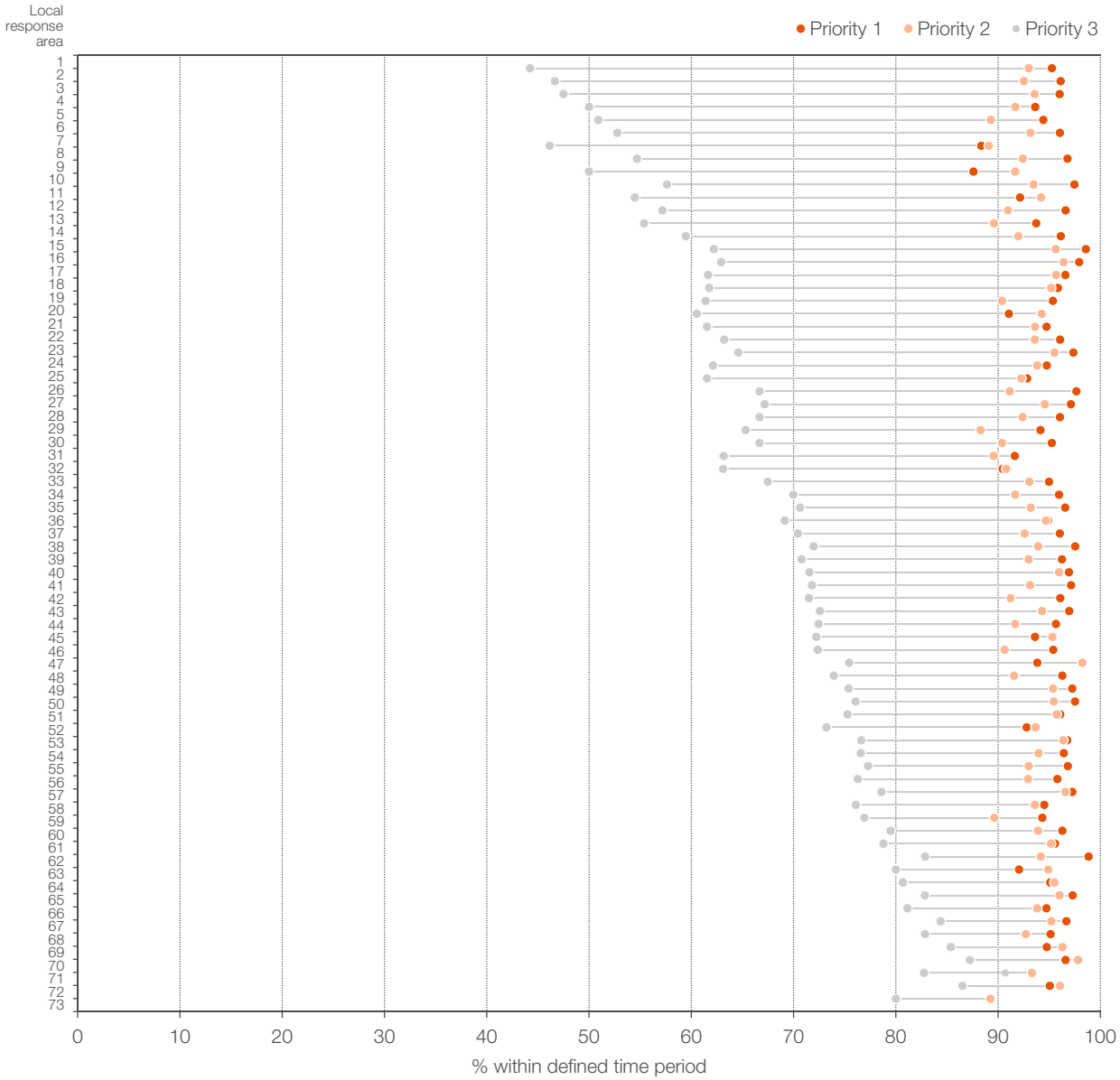


Figure 44 24-hour (on call) local response area call to ambulance arrival within defined times, by priority category, NSW, July to September 2016

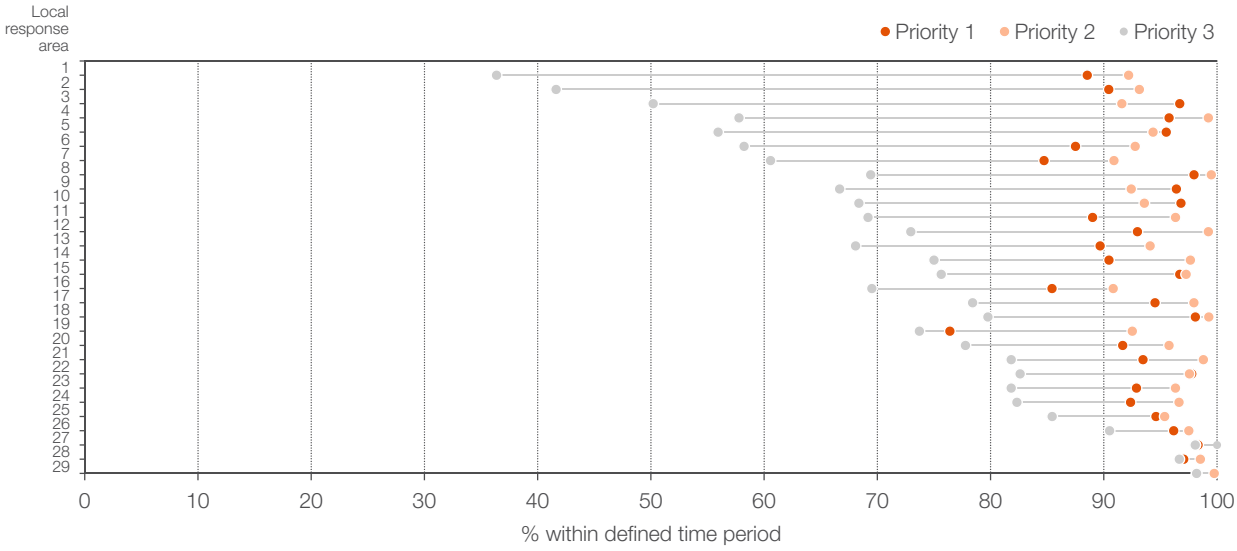
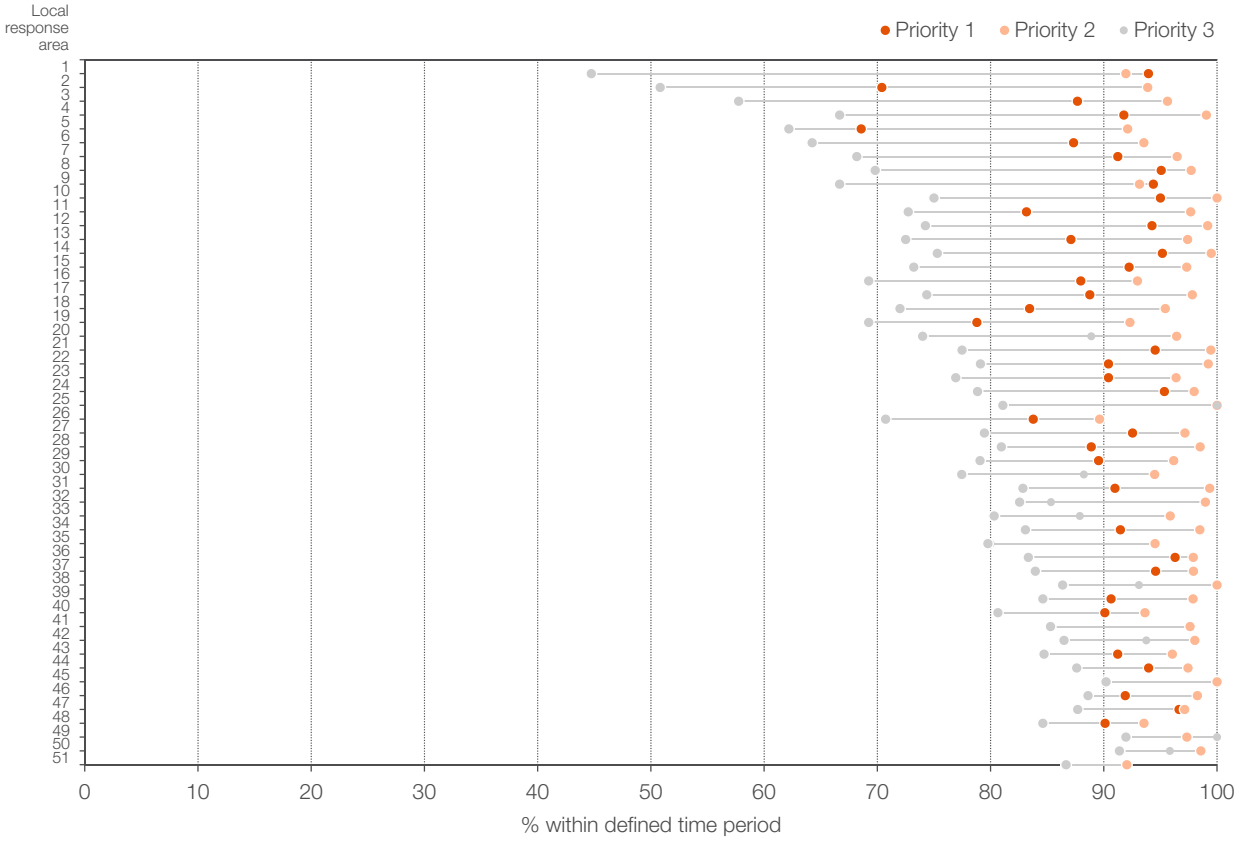


Figure 45 Non-24-hour (on call) local response area call to ambulance arrival within recommended times, by priority category, NSW, July to September 2016



Supression rules for meaningful reporting

Suppression rules aim to limit reporting to units for which performance information is fair and meaningful. A series of criteria were developed to identify LRAs with small or unstable numbers of responses.

These criteria exclude LRAs with:

- less than five consecutive quarters of data prior to July to September 2016
- on average number of responses per quarter that was less than 100
- a coefficient of variation above 10%
- volunteer and community LRA type

These exclusions are not mutually exclusive. Some LRAs met multiple exclusion criteria – for example, 31 were excluded because they had on average fewer than 100 responses per quarter, they had a coefficient of variation above 10%, and they were a volunteer and community LRA type. A total of 113 LRAs were excluded, leaving 153 in the LRA-level analyses (Figure 46).

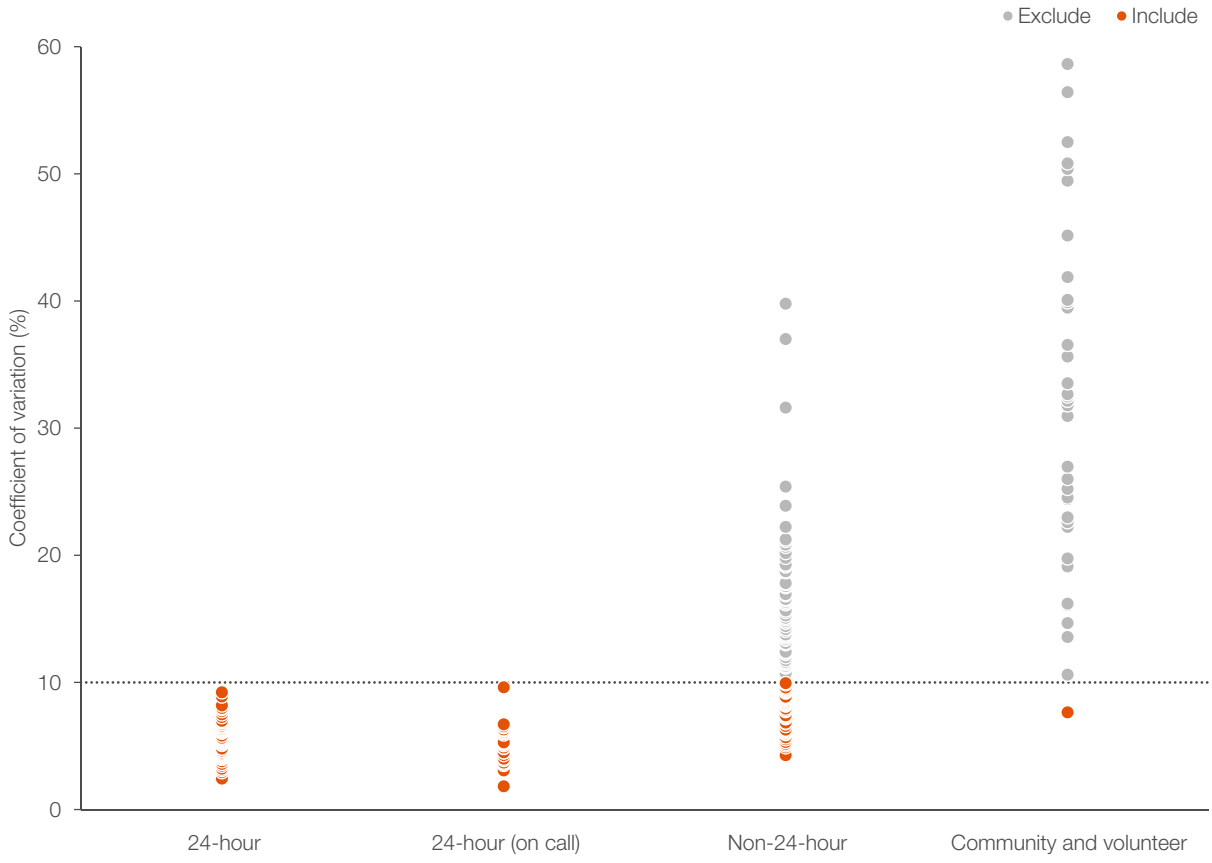
Figure 46 Application of suppression rules, local response areas, NSW, July to September 2016

Number of LRAs prior to exclusions	266
LRAs without five consecutive quarters of data	7
LRAs with an average < 100 responses per quarter	1
LRAs with a coefficient of variation > 10%	55
LRA type (volunteer and community)	1
LRAs with an average < 100 responses per quarter and a coefficient of variation > 10%	14
LRAs with a coefficient of variation > 10% and LRA type (volunteer and community)	4
LRAs with an average < 100 responses per quarter and a coefficient of variation > 10% and LRA type (volunteer and community)	31
Total number of LRAs included for quarterly reporting	**153**

After suppression, 55% of LRAs remained in the analyses. The majority of suppressions occurred in the non-24-hour and volunteer and community LRA types (Figure 47).

Suppressed data for individual LRAs are not featured in graphs depicting LRA-level performance, however they are included in measures of activity and performance at aggregated levels of reporting such as zone and NSW.

Figure 47 Local response areas by coefficient of variation in the number of responses, NSW, July 2014 to September 2016



The impact of applying the suppression rules

To illustrate the impact of the suppression rules (i.e. less than 100 responses per quarter, coefficient of variation of 10% or more, volunteer and community LRAs), analyses were run with and without application of the rules (Figures 48 and 49).

Overall, 113 LRAs had their results suppressed and the distribution ranged from 0 to 20 across the zones (Figure 44).

Levels of random variation were also assessed by calculating confidence intervals around the median response time. The instability of estimates for non-24-hour stations is apparent – even for those which met the reporting criteria – further supporting the proposal not to nominally report at a local response area level (Figure 49).

Figure 48 The effect of suppression rules, zone and local response area performance with no suppression rules applied, median response time, NSW, July to September 2016

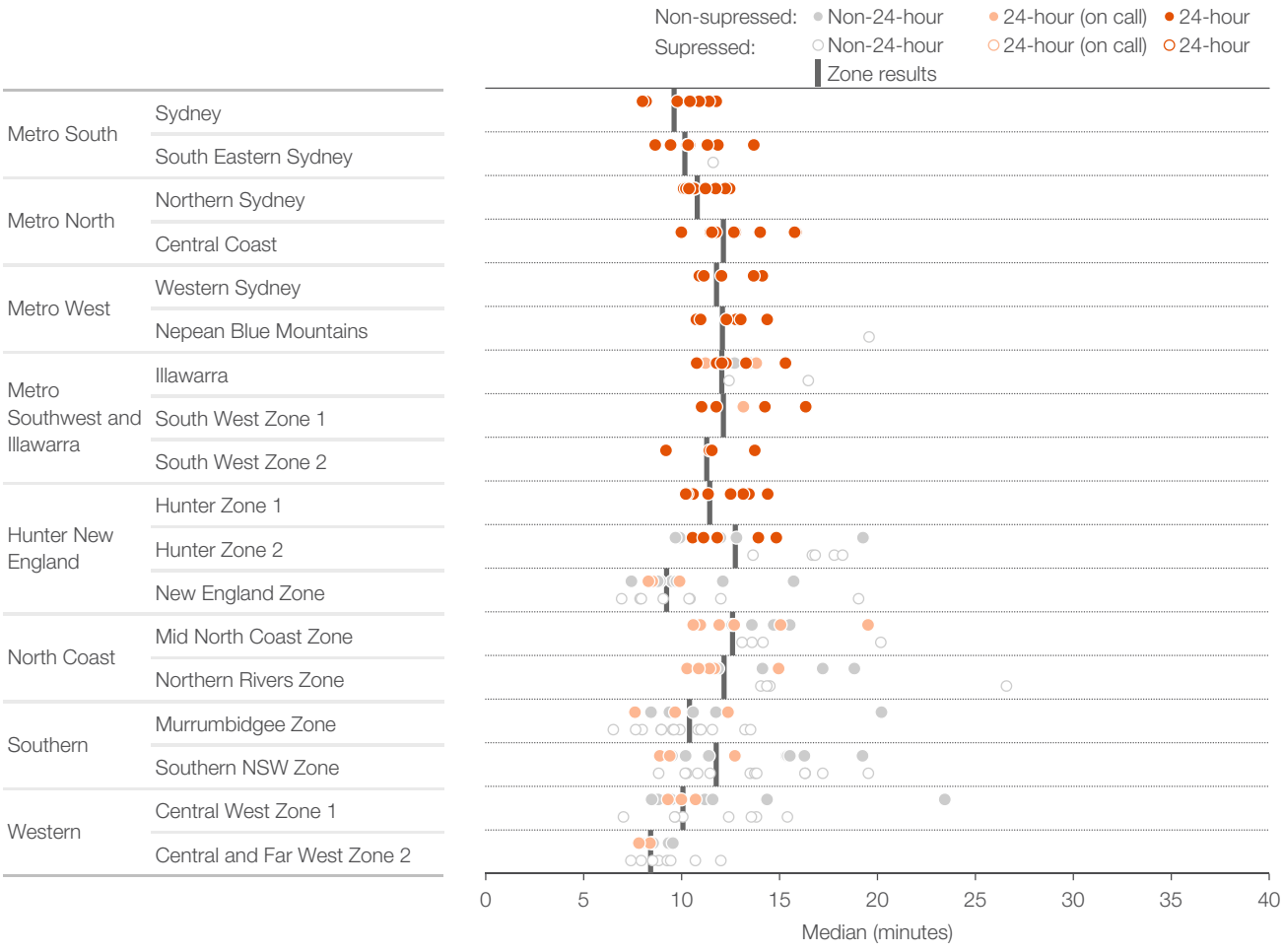
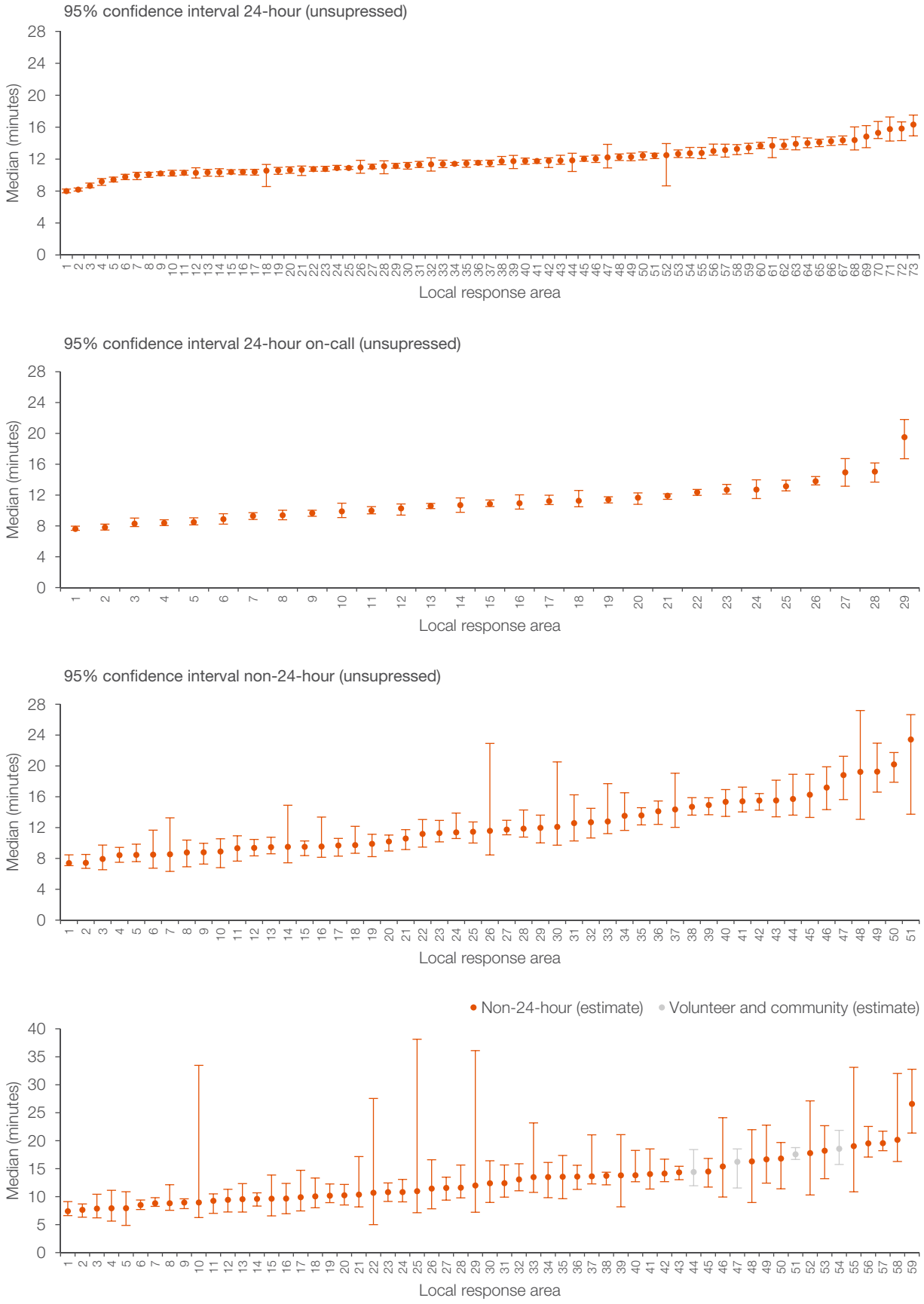


Figure 49 Assessing variation of median response time using 95% confidence intervals, by local response area station type, NSW, July to September 2016



Sensitivity of measures

Priority 1 thresholds

Time-based benchmarks are often defined alongside a threshold performance level. Reporting of LRA performance focuses on whether or not a threshold percentage of responses achieved the defined benchmarks for the priority categories. Two time periods were considered: first, priority 1 call to ambulance arrival times within 15 minutes (Figures 50 and 51), and second, priority 1 call to ambulance

Figure 50 Priority 1 call to ambulance arrival times within 15 minutes, 50%, 60%, 70%, 80% and 90% thresholds, zones and local response areas, July to September 2016

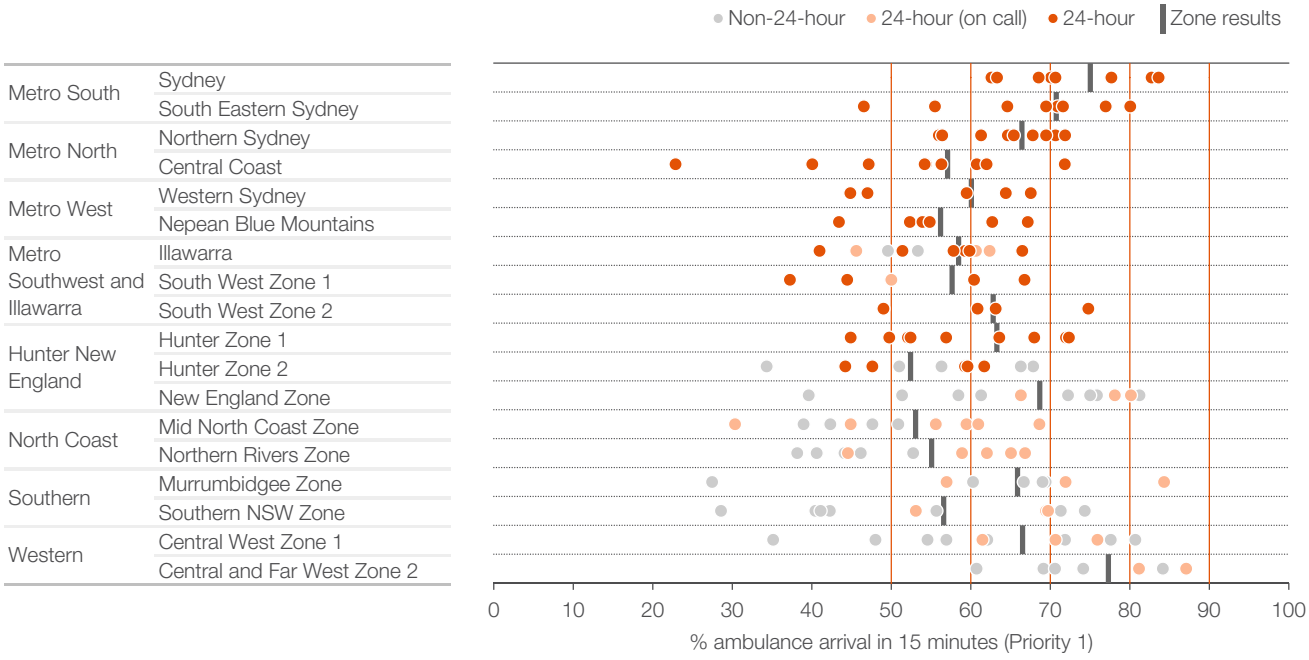
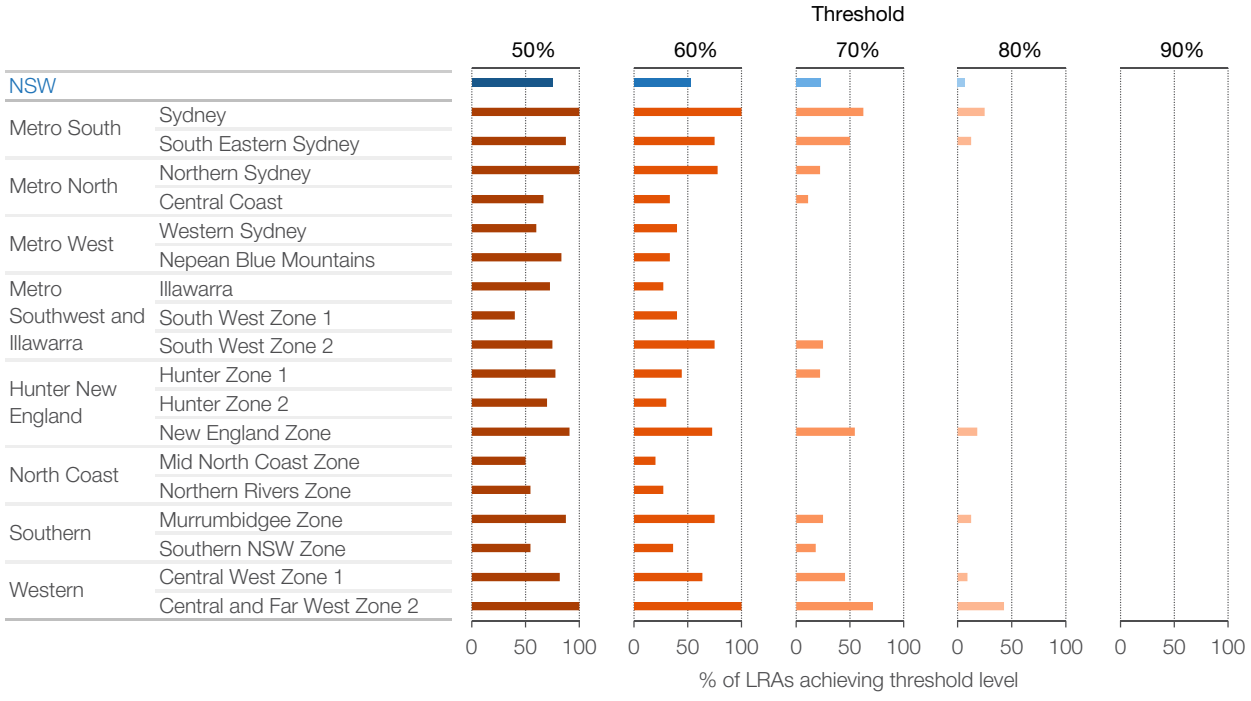


Figure 51 Percentage of local response areas that achieved 15 minute benchmark for priority 1 responses, 50%, 60%, 70%, 80% and 90% thresholds, NSW, July to September 2016



arrival times within 30 minutes (Figures 52 and 53). Thresholds of 50%, 60%, 70%, 80% and 90% were assessed. At a NSW level, 62.5% of priority 1 call to ambulance times were within 15 minutes, however no LRAs reached the 90% threshold. *Healthcare Quarterly* reporting will focus on the 90% threshold for the benchmark time of 30 minutes.

Figure 52 **Priority 1 call to ambulance arrival times within 30 minutes, 50%, 60%, 70%, 80% and 90% thresholds, zones and local response areas, July to September 2016**

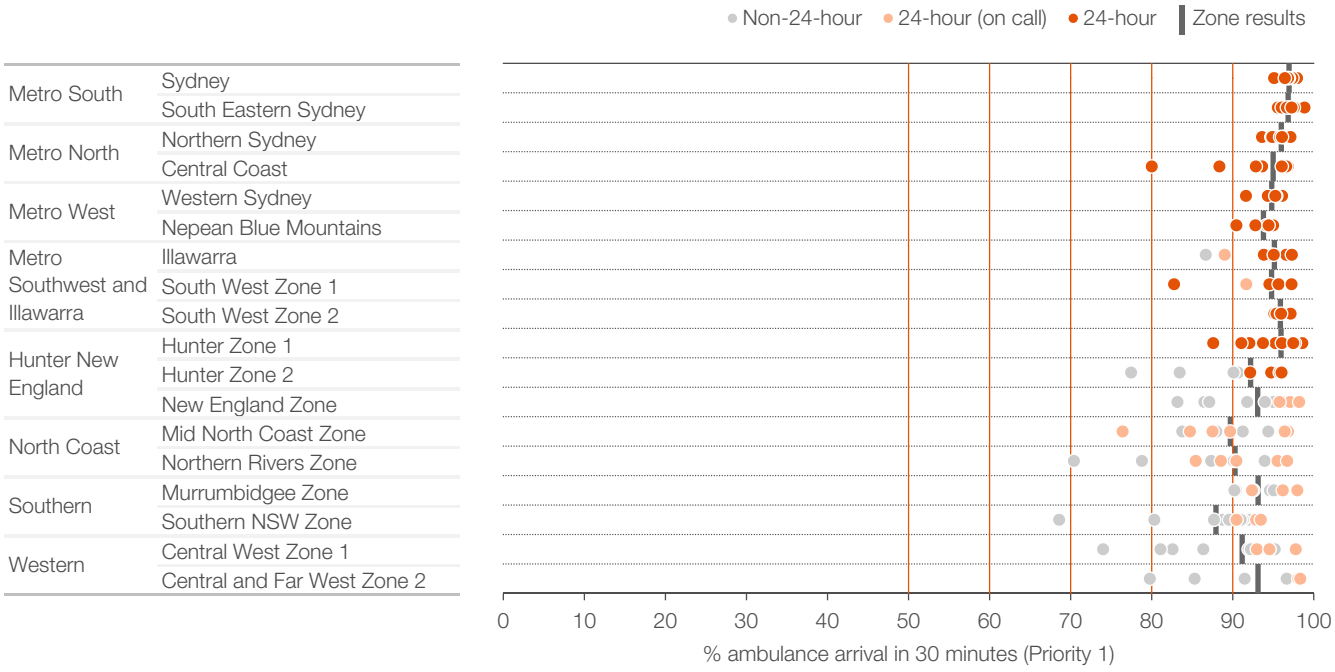
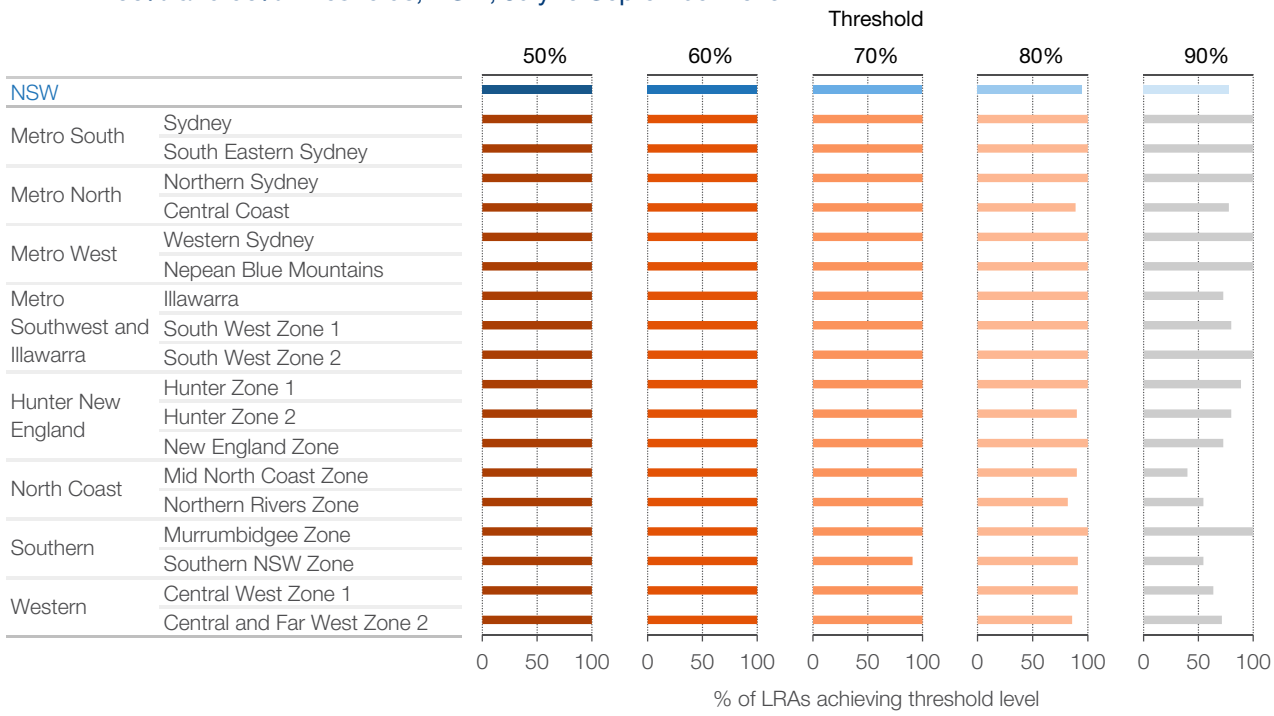


Figure 53 **Percentage of LRAs that achieved 30 minute benchmark for priority 1 responses, 50%, 60%, 70%, 80% and 90% thresholds, NSW, July to September 2016**



Sensitivity of measures

Priority 2 thresholds

Two time periods were also considered for priority 2 responses: first, priority 2 call to ambulance arrival times within 30 minutes (Figures 54 and 55), and second, priority 2 call to ambulance arrival times within 60 minutes (Figures 56 and 57). Thresholds of 50%, 60%, 70%, 80% and 90% were assessed. At a NSW level, 73.8% of priority 2 call to ambulance times were within 30 minutes. When the threshold

Figure 54 Priority 2 call to ambulance arrival times within 30 minutes, 50%, 60%, 70%, 80% and 90% thresholds, zones and local response areas, July to September 2016

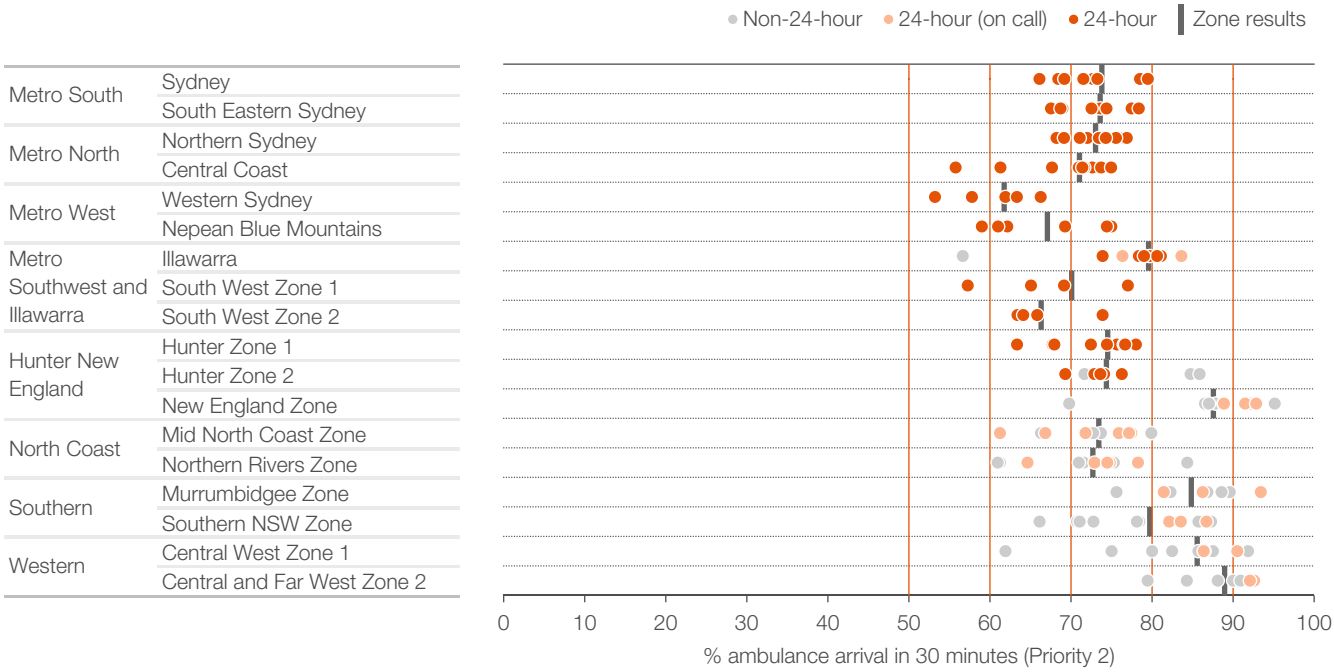
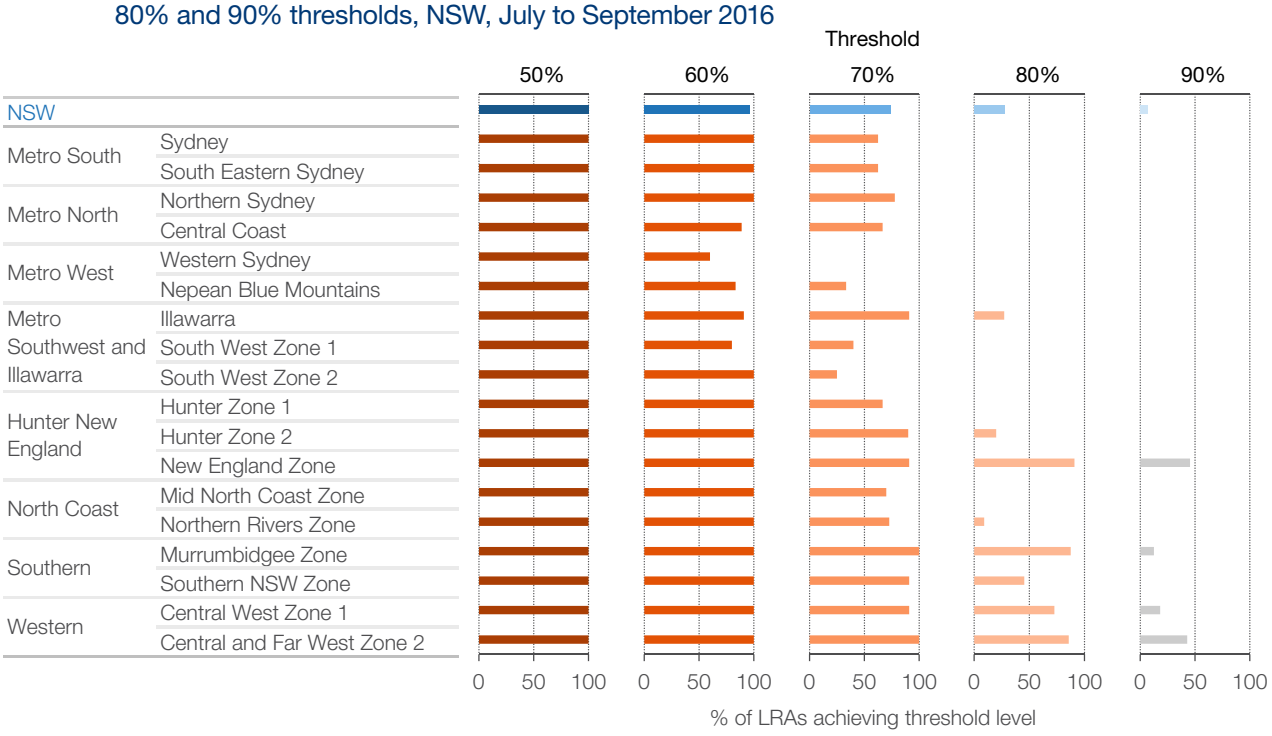


Figure 55 Percentage of LRAs that achieved 30 minute benchmark for priority 2 responses, 50%, 60%, 70%, 80% and 90% thresholds, NSW, July to September 2016



was set at 90%, four zones achieved the benchmark. In comparison, 94.1% of NSW priority 2 call to ambulance arrival times were within 60 minutes and almost all local response areas exceeded the 90% achievement threshold (Figures 56 and 57). *Healthcare Quarterly* reporting will focus on the 90% threshold for the benchmark time of 60 minutes.

Figure 56 Priority 2 call to ambulance arrival times within 60 minutes, 50%, 60%, 70%, 80% and 90% thresholds, zones and local response areas, July to September 2016

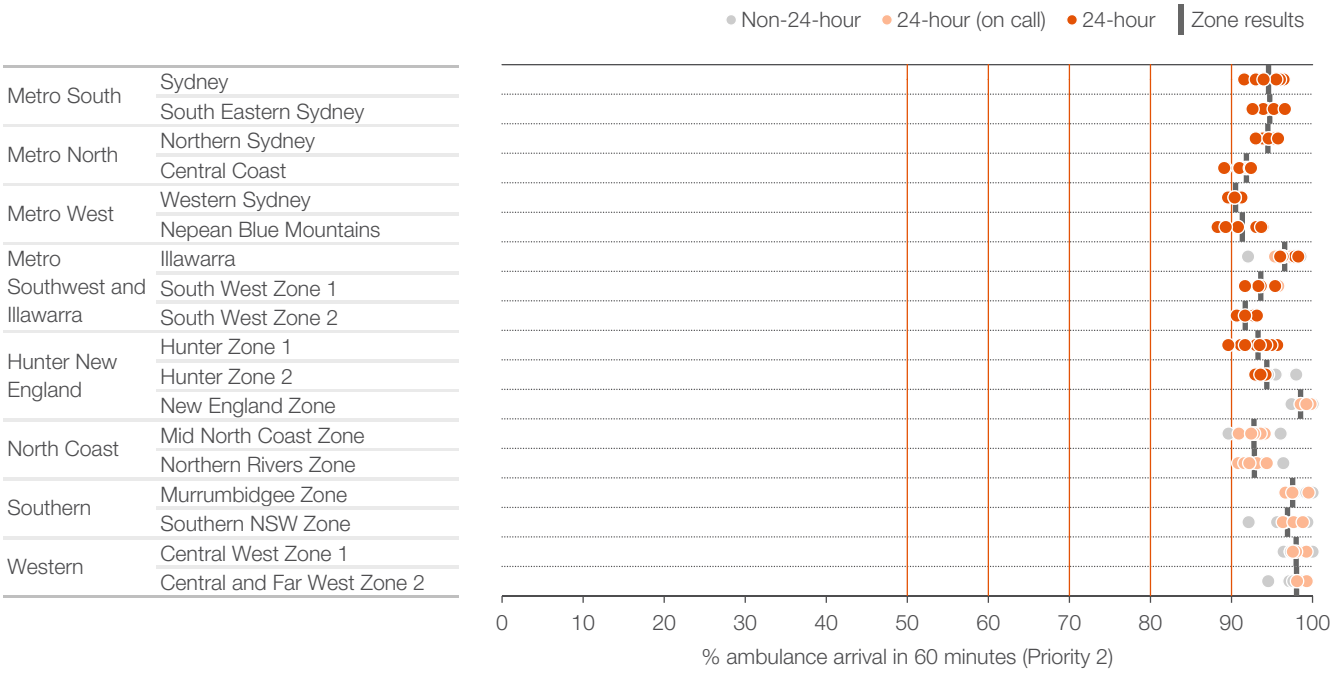
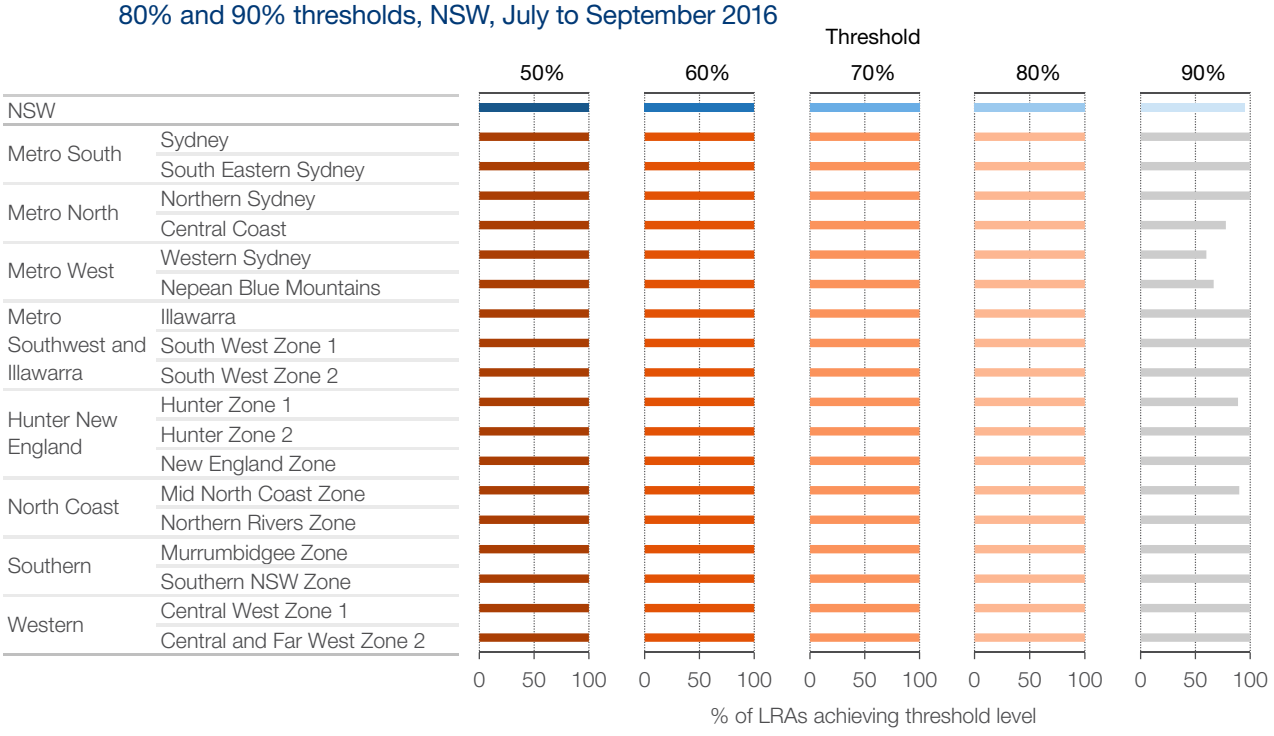


Figure 57 Percentage of LRAs that achieved 60 minute benchmark for priority 2 responses, 50%, 60%, 70%, 80% and 90% thresholds, NSW, July to September 2016



Sensitivity of measures

Propensity to transport

A potential confounder to be considered in ambulance performance reporting is whether there are geographically or organisationally-based differences in the propensity to transport patients.

At an LRA level, there was substantial variation in the percentage of incidents that resulted in patient transport. For 24 hour local response areas in the July to September 2016 quarter, this percentage ranged from 57.4% to 79.9%; for 24 hour (on-call) it ranged from 59.8% to 82.7%; and for non-24 hour local response areas it ranged from 54.4% to 89.9% (Figure 55).

At a zone level the percentage of incidents which resulted in a patient transport were more consistent, ranging from 65% in Hunter Zone 1 to 78% in Central and Far West (Figure 59).

When stratified into priority categories, differences across zones were less pronounced (Figure 60).

Figure 58 Percentage of incidents that resulted in patient transport, by local response area type, NSW, July to September 2016

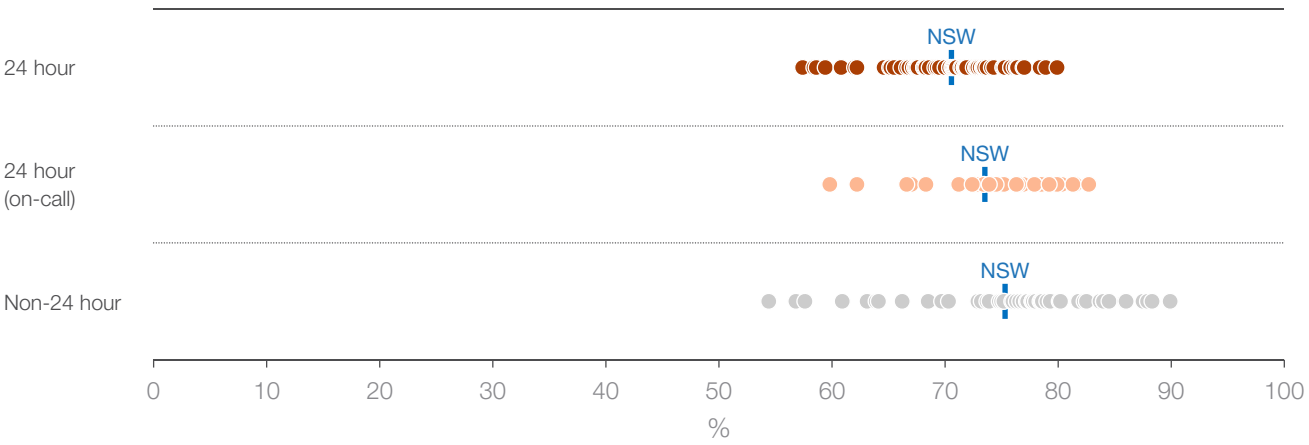


Figure 59 Percentage of incidents that resulted in patient transport, by zone, NSW, July to September 2016

Metro South	Sydney
	South Eastern Sydney
Metro North	Northern Sydney
	Central Coast
Metro West	Western Sydney
	Nepean Blue Mountains
Metro Southwest and Illawarra	Illawarra
	South West Zone 1
	South West Zone 2
Hunter New England	Hunter Zone 1
	Hunter Zone 2
	New England Zone
North Coast	Mid North Coast Zone
	Northern Rivers Zone
Southern	Murrumbidgee Zone
	Southern NSW Zone
Western	Central West Zone 1
	Central and Far West Zone 2

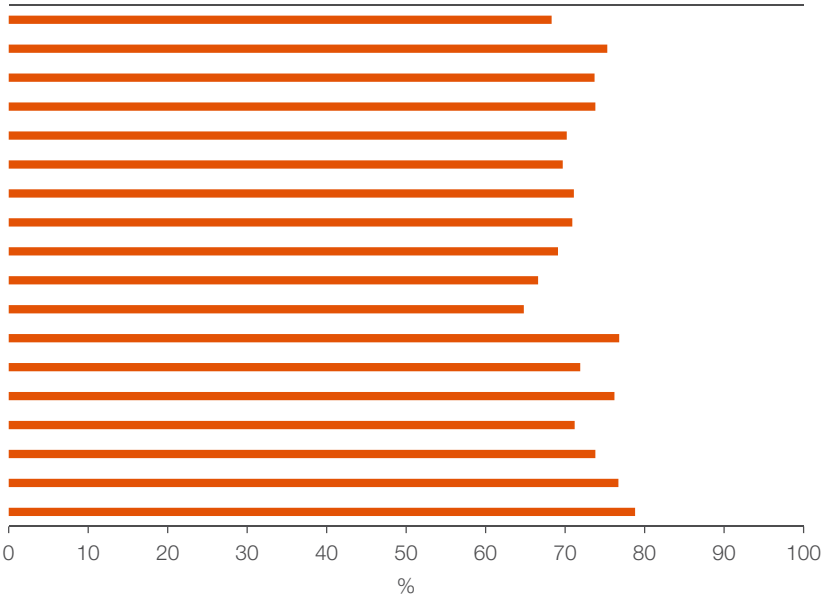
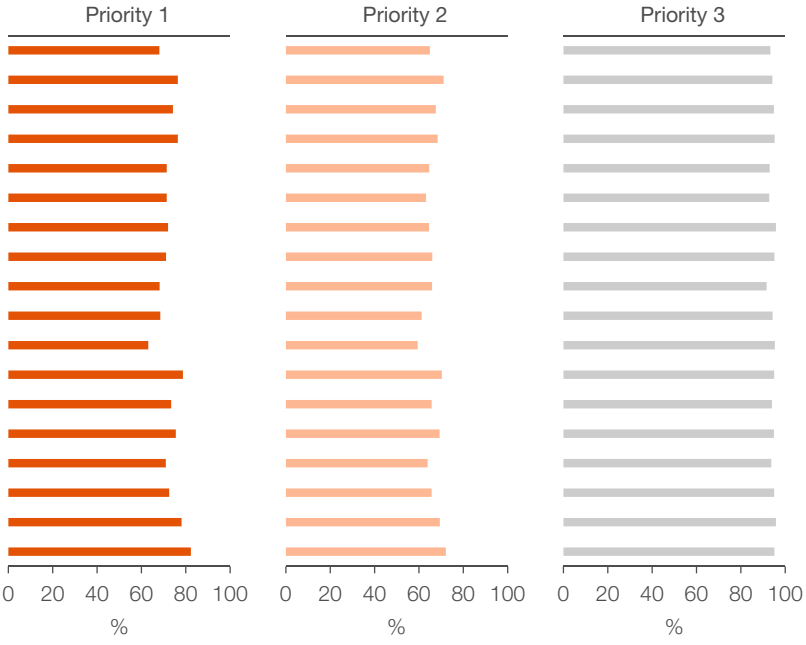


Figure 60 Percentage of incidents that resulted in patient transport, by priority category and zone, NSW, July to September 2016

Metro South	Sydney
	South Eastern Sydney
Metro North	Northern Sydney
	Central Coast
Metro West	Western Sydney
	Nepean Blue Mountains
Metro Southwest and Illawarra	Illawarra
	South West Zone 1
	South West Zone 2
Hunter New England	Hunter Zone 1
	Hunter Zone 2
	New England Zone
North Coast	Mid North Coast Zone
	Northern Rivers Zone
Southern	Murrumbidgee Zone
	Southern NSW Zone
Western	Central West Zone 1
	Central and Far West Zone 2



Attribution of performance at a zone level

Call to ambulance arrival time outside defined period

While there is insufficient stability to publicly report LRA performance in call to ambulance arrival times (see pages 43-46), it is important to describe and monitor where patients wait longer than the defined time periods and to reflect on system performance.

Overall in NSW for the July to September 2016 quarter, 91.9% call to ambulance arrival times were within the defined period. A series of analyses were undertaken to describe the call to ambulance arrival times that were outside the defined limits (Figure 61-63).

In terms of LRA type, of the call to ambulance arrival times that were above 30 minutes for priority 1, 24-hour stations accounted for 54.6%; 24-hour with on call stations accounted for 18.4%; and non-24-hour stations accounted for 24.6% (Figure 62).

On a geographical basis, Mid-North Coast accounted for the largest share or call to ambulance arrival times that were longer than 30 minutes (9.3%) (Figure 63).

For priority 2, Western Sydney (12.7%) and Northern Rivers (13.3%) accounted for the largest share of call to ambulance arrival times that were longer than 60 minutes [data not shown].

Figure 61 Number of responses and call to ambulance arrival times outside defined times, responses by zone and priority category

	Priority 1		Priority 2	
	Number of priority 1 responses with a call to ambulance arrival time	Number of priority 1 call to ambulance arrival times over 30min	Number of priority 2 responses with a call to ambulance arrival time	Number of priority 2 call to ambulance arrival times over 60min
Sydney	9751	298	9115	495
South Eastern Sydney	6964	220	7059	373
Northern Sydney	6940	278	7093	391
Central Coast	5801	291	5919	483
Western Sydney	8905	462	7399	703
Nepean Blue Mountains	5988	372	5241	454
Illawarra	5947	289	6132	211
South West Zone 1	5563	290	4726	302
South West Zone 2	6922	285	5611	466
Hunter Zone 1	5664	229	5975	402
Hunter Zone 2	3386	264	3421	193
New England Zone	2442	169	2839	42
Mid North Coast Zone	4878	502	4938	357
Northern Rivers Zone	4302	418	4443	319
Murrumbidgee Zone	3449	237	3866	95
Southern NSW Zone	3264	394	3560	110
Central West Zone 1	2738	242	3015	61
Central and Far West Zone 2	2005	138	2227	44

Figure 62 Distribution of NSW priority 1 call to ambulance arrival times over 30 minutes, by local response area type, July to September 2016

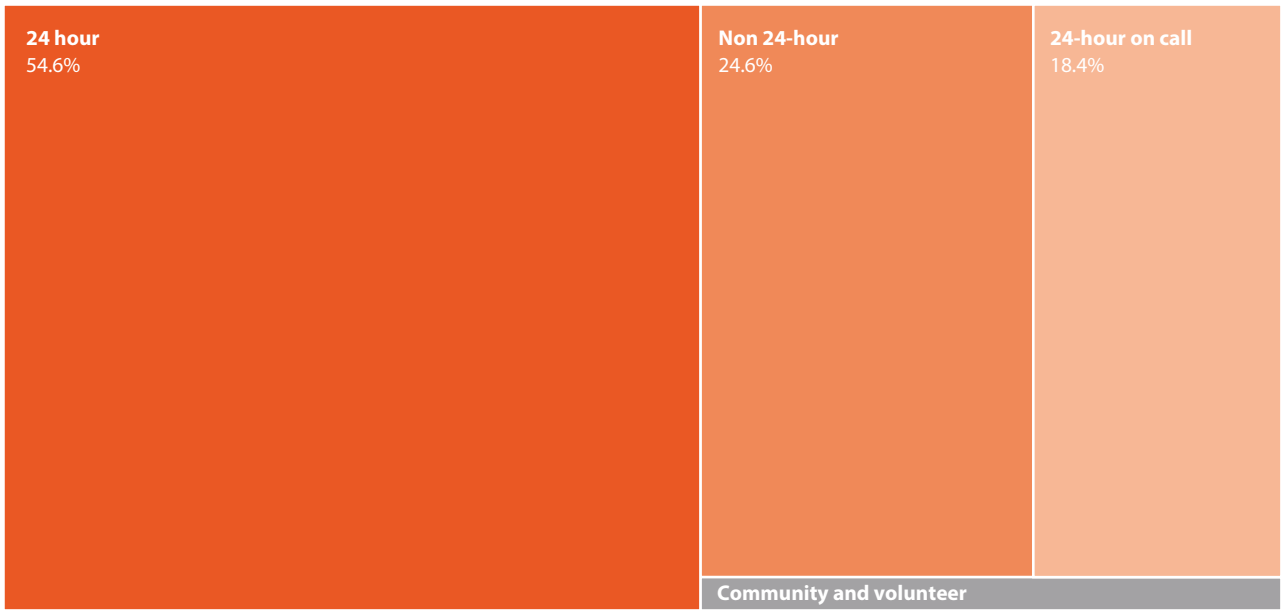
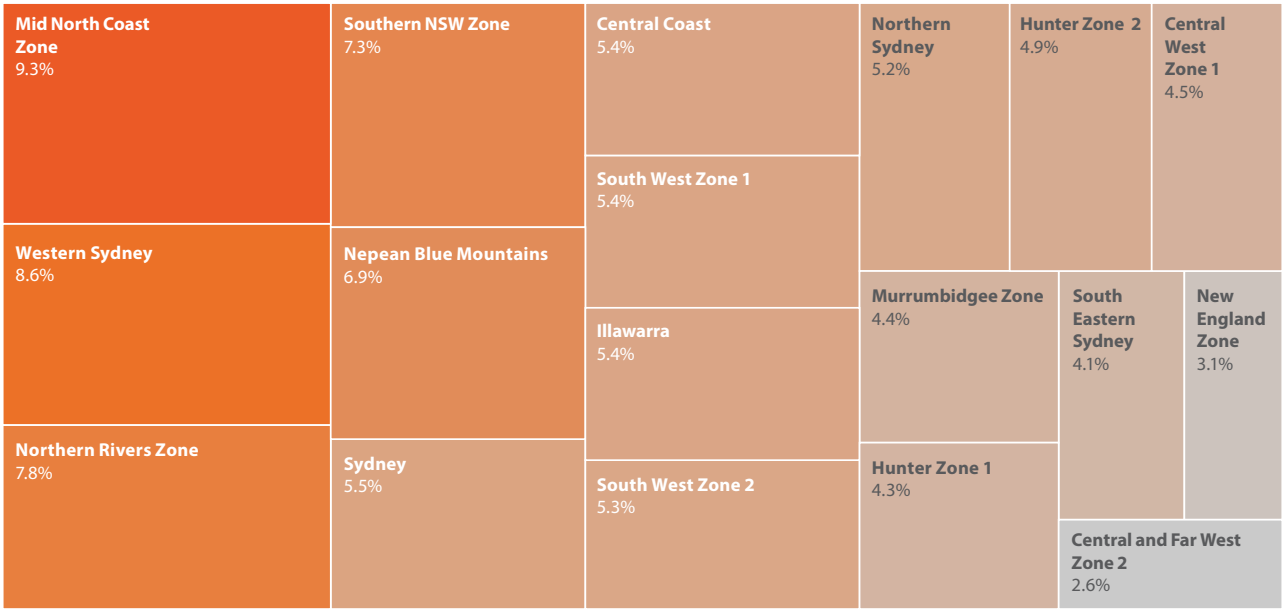


Figure 63 Distribution of NSW priority 1 call to ambulance arrival times over 30 minutes, by zone, NSW, July to September 2016



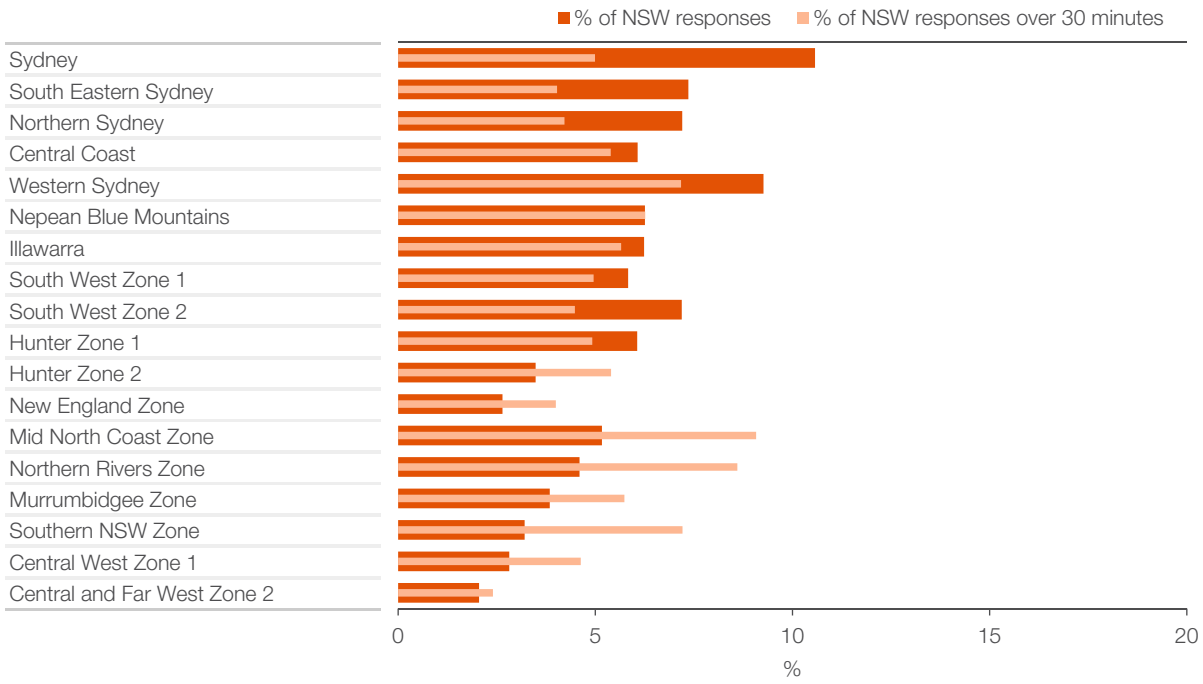
Reflecting on system-level performance

Public reporting on a quarterly basis will describe the characteristics of patients for whom the defined call to ambulance arrival times were not achieved.

While LRA results will not be reported individually, the proportion of responses across all quarters that were outside call to ambulance arrival time benchmarks and that occurred in each zone will be published. To help inform interpretation, this information will be presented alongside the proportion of all responses that occurred in each zone (Figures 64-65).

Relative to the percentage of priority 1 responses in each zone, non-metropolitan zones contribute disproportionately more to the total number of call to ambulance arrival times in NSW that are over 30 minutes (Figure 64).

Figure 64 **Distribution of priority 1 responses in NSW and percentage of call to ambulance arrival times over 30 minutes, by zone, NSW, July to September 2016**



The zone that contributed the fewest priority 1 responses above the 30 minute target to the NSW total was Central and Far West (2.5% of the NSW total). The zone that contributed the most was Mid North Coast (9.3% of all NSW responses above 30 minutes). Notably, that zone had a comparatively smaller proportion of the state's priority 1 responses (5.1% of responses) (Figure 64).

For priority 2, Western Sydney had a disproportionately large number of responses exceeding the defined time, relative to its share of responses in NSW (Figures 65).

Figure 65 Distribution of priority 2 responses in NSW and the percentage of call to ambulance arrival times over 60 minutes, by zone, NSW, July to September 2016

Assessing performance in the context of activity

Surge day analysis

Mobile emergency health services can face unexpected fluctuations in demand – whether this be from coincident events, extreme weather or natural disasters. The resilience of the service to respond to fluctuating need is fundamental to a high-performing ambulance system.

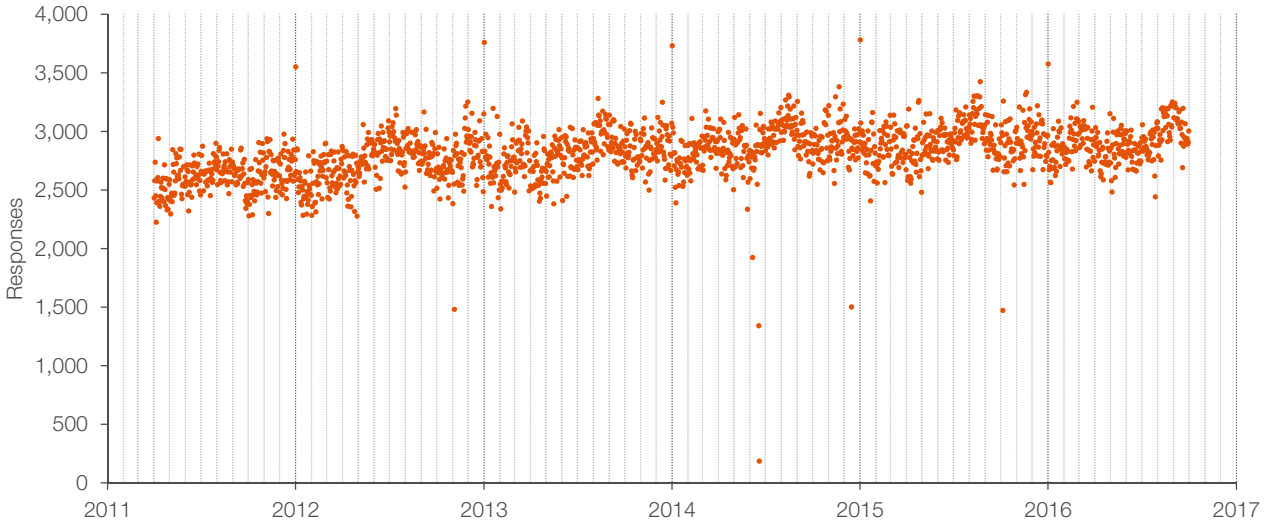
Time series data spanning five years from 2011 to 2016 was used to predict the daily number of ambulance responses while accounting for trends, and seasonal and other cyclical patterns that were present in the data (Figure 66).

A first order autoregression model was fit to correct for the serial correlation of the error terms between time points. The daily number of ambulance responses was modelled as a function of a series of indicator variables that controlled for long term trends and seasonality: year, month of the year, week of the year, and day of the week. An indicator variable was also included to denote a policy change in the transportation of non-urgent cases in the autumn of 2014. Prediction intervals were used to determine the range within which the expected number of ambulance responses on a single day was likely to fall. Different prediction interval levels were specified to assess the face validity of the surge days identified using the model. More conservative prediction intervals flagged fewer surge days.

A key consideration in selecting the prediction interval level was striking a balance between being too sensitive to day-to-day fluctuations in the observed number of responses (i.e. over-identifying surge days in the quarter), and not being sensitive enough to genuine fluctuations in the data (i.e. under-identifying surge days). The prediction interval limit for the analysis was set at 75% after having assessed less conservative (50%) and more conservative (90%) prediction intervals on July to September 2016 data.

Days above the upper limit were considered to have a significantly greater than expected number of ambulance responses (i.e. a surge day) (Figure 67). Conversely, days below the lower limit were deemed to have a significantly lower than expected number of ambulance responses (i.e. a lull day). Relative to the 75% prediction interval, the 90% prediction interval identified a single surge day in September. Conversely, the 50% prediction interval flagged 16 surge days over the quarter [data not shown].

Figure 66 Time series of daily number of responses, NSW, 2011 to 2016



To highlight times of ambulance service resilience, analyses were conducted to explore the relationship between ambulance activity and performance. For each day in the July to September 2016 quarter, the results from the surge analysis were presented alongside performance data (Figure 68).

Figure 67 Days in the quarter with an above or below expected number of responses (priority 1 to priority 3), 70% prediction interval, NSW, July to September 2016

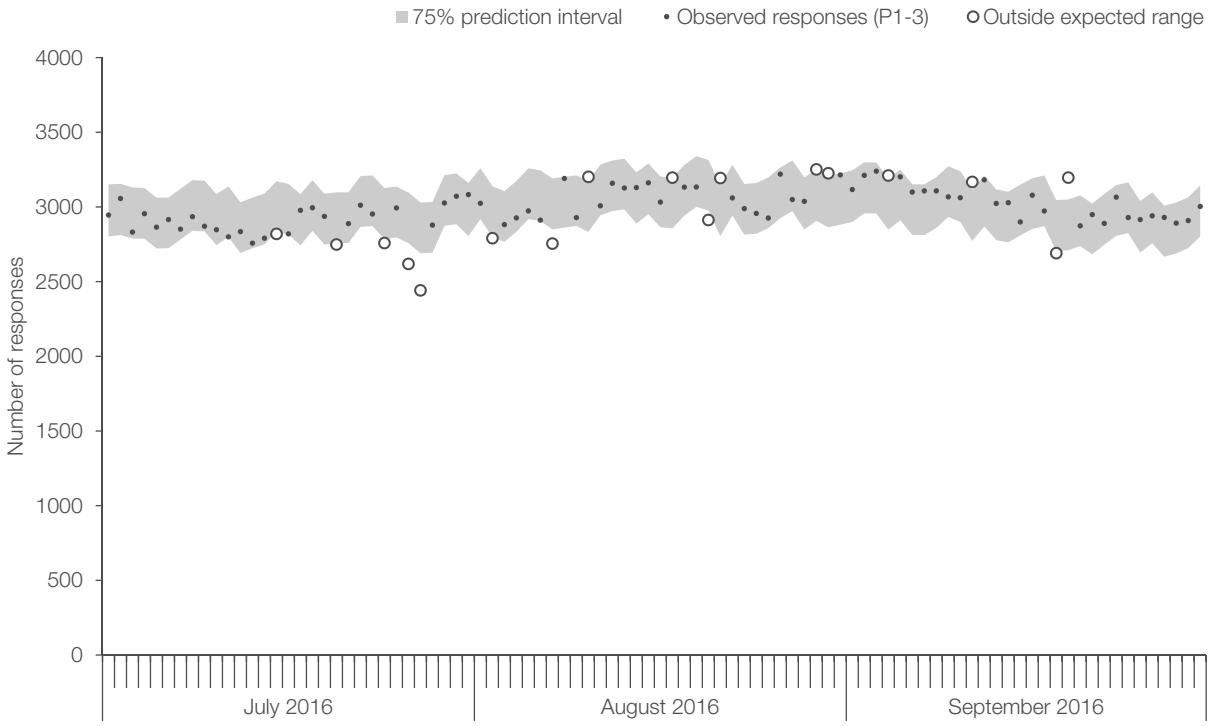
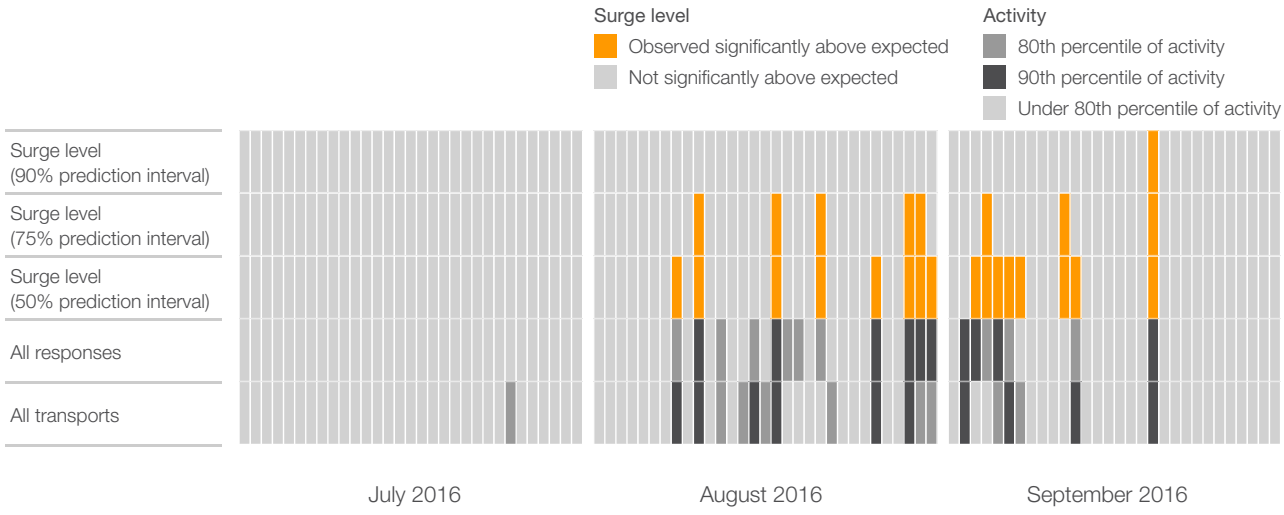


Figure 68 Days in the quarter with an above or below expected number of responses (priority 1 to priority 3), and relative performance measures, NSW, July to September 2016



Assessing integrated performance

Turnaround time from two perspectives

Turnaround time is affected by ambulance crews, emergency department staff and the interactions between them.

In the January to March 2017 quarter, priority 1 and 2 turnaround time in NSW was 34.6 minutes and ranged from 21.9 minutes in Southern NSW to 42.1 in Western Sydney (Figure 69).

While most LRAs have a single hospital to which they take most patients, there are around 30, mostly metropolitan LRAs, that are 'less exclusive' and take patients to a number of different hospitals.

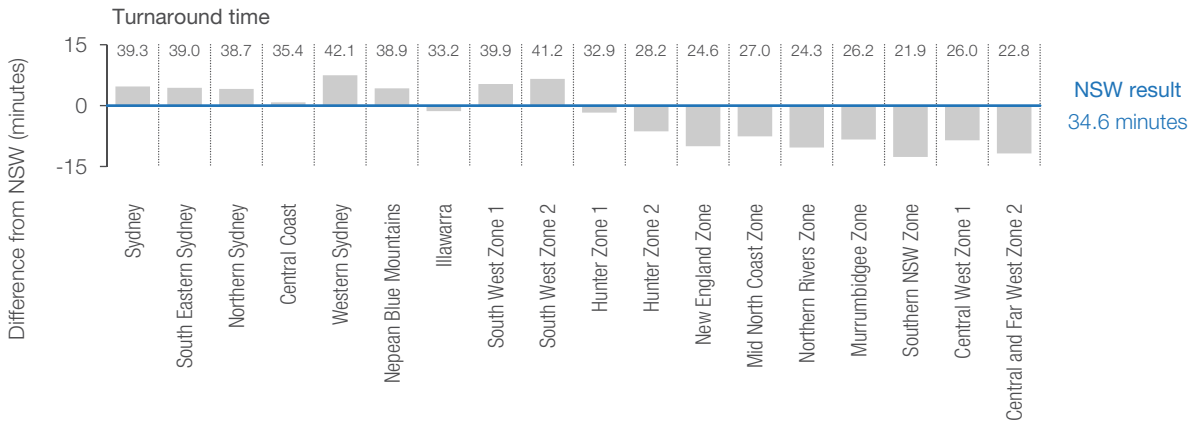
Similarly, while many hospitals receive patients almost exclusively from a single ambulance LRA, there are about 20 'less exclusive' hospitals that are served by ambulance crews from different LRAs.

In the January to March 2017 quarter, Westmead Hospital received patients from ambulances based in nine different LRAs. The median turnaround times for crews from the different LRAs varied from 41.8 minutes to 53.6 minutes (11.8 minute range) (Figure 70).

Auburn was one of the local response areas that took patients to Westmead Hospital. Altogether, crews based in Auburn LRA took patients to eight different hospitals. The median turnaround times for the Auburn-based crews varied across hospitals from 35.7 to 45.3 minutes (9.6 minute range) (Figure 71).

Less exclusive LRAs are defined as those that take patients to the same hospital less than 70% of the time; and 'less exclusive' hospitals are defined as those that receive patients from the same LRA less than 70% of the time.

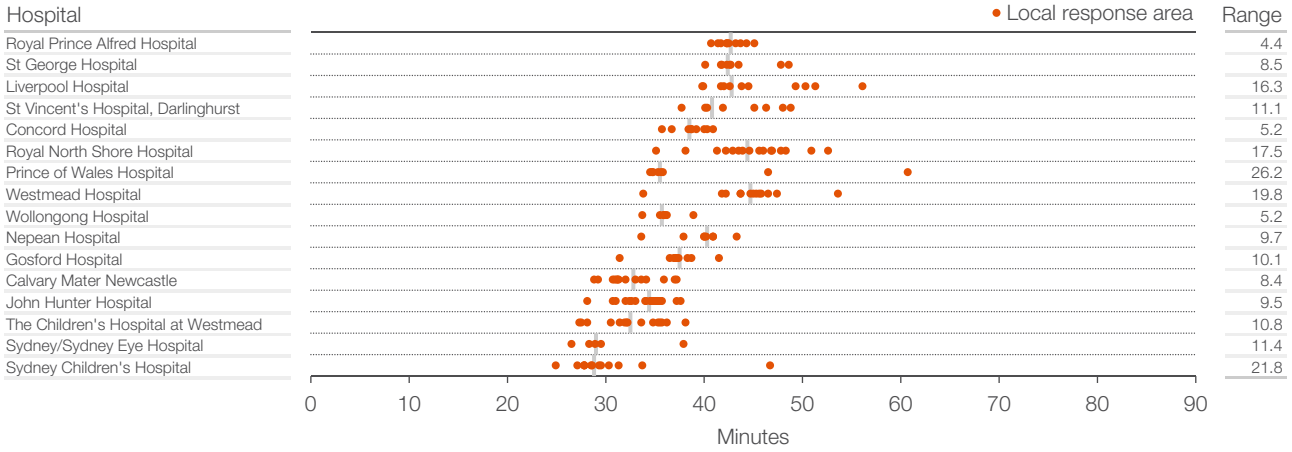
Figure 69 Median turnaround time, by zone, priority 1 and 2, relative to NSW, January to March 2017



In general, there was greater variability within each LRA's relationships (spread of results ranged from 21.3 to 53.6 minutes) than within each hospital's relationships (spread of results ranged from 24.9 to 60.7 minutes).

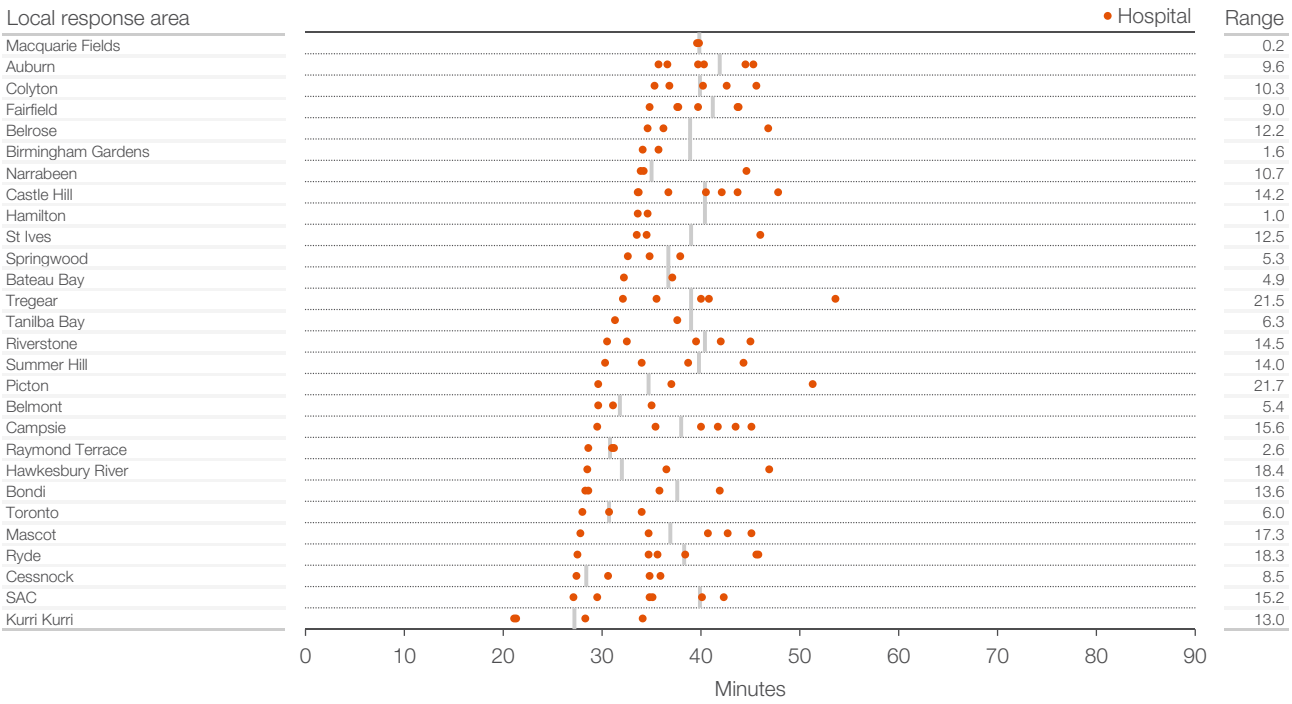
While both ambulance crews and hospital emergency departments influence turnaround times, it appears that hospitals – and in particular larger hospitals – have a greater influence on results (Figures 70 and 71).

Figure 70 Median turnaround time for priority 1-2 responses, hospital perspective (peer group A), January to March 2017



Note: This graph shows median turnaround times for peer group A hospitals that receive patients from the same local response area less than 70% of the time.

Figure 71 Median turnaround time for priority 1-2 responses, local response area perspective (24-hour), January to March 2017



References

1. NSW Health. Strategic Review of the Ambulance Service of NSW Final Report. June 2012. http://www.health.nsw.gov.au/about/nswhealth/Pages/asnsw_review.aspx
2. 2014-2015 Service Agreement – An agreement between Secretary NSW Health and Ambulance Service NSW for the period 2014-2015. Available at http://www.ambulance.nsw.gov.au/Our-performance/Governance-and-Accountability.html
3. Australian Government Productivity Commission Report on Emergency management, 2015. Available at http://www.pc.gov.au/research/recurring/report-on-government-services/2015/emergency-management
4. NSW Ambulance, Our performance 2015-16. Available at http://www.ambulance.nsw.gov.au/Our-performance.html
5. Council of Ambulance Authorities Data Dictionary – v1 – July 2016.
6. Ambulance Victoria's Performance 2016/17 Quarter 2 October to December 2016. Available at http://ambulance.vic.gov.au/about-us/our-performance/
7. Queensland Ambulance Service, Public Performance Indicators financial year first quarter - July to September 2016. Available at https://www.ambulance.qld.gov.au/index.html
8. NSW Ambulance CAD Data Model Documentation Business Rules and Data Dictionary. Version 2.1 April 2016.
9. NHS England Ambulance System Indicators 2016 Available at https://www.england.nhs.uk/statistics/statistical-work-areas/ambulance-quality-indicators/ambulance-quality-indicators-data-2016-17/
10. Ontario Land Ambulance Response Time Standard 2016. Available at http://www.health.gov.on.ca/english/public/Program/ehs/land/responsetime.html

Acknowledgements

The Bureau of Health Information (BHI) is the main source of information for the people of NSW about the performance of their public healthcare system. A NSW board-governed organisation, BHI is led by Chairperson Professor Carol Pollock and Acting Chief Executive Dr Kim Sutherland.

We would also like to acknowledge the contribution of Dr Jean-Frederic Levesque for his leadership and guidance at BHI during the development of this report. We would like to thank our expert advisors, colleagues at the Ministry of Health and NSW Ambulance, reviewers and staff who contributed to the report.

External Advisors and Reviewers

NSW Ambulance

Queensland Ambulance Service

Ambulance Victoria

Health Quality and Safety Commission New Zealand

Bureau of Health Information Project Team

Analytics

Renee Carter

Carolynn Fredericks

Jill Kaldor

Behnoosh Khalaj

Research

Lilian Daly

Design

Ed Bury

Adam Myatt

Efren Sampaga

Mark Williams

Communications and Stakeholder Engagement

Rohan Lindeman

Karen Perini

Stephanie Watson

bureau of health information

About the Bureau of Health Information

The Bureau of Health Information (BHI) is a board-governed organisation that provides independent information about the performance of the NSW public healthcare system.

BHI was established in 2009 to provide system-wide support through transparent reporting.

BHI supports the accountability of the healthcare system by providing regular and detailed information to the community, government and healthcare professionals. This in turn supports quality improvement by highlighting how well the healthcare system is functioning and where there are opportunities to improve.

BHI manages the NSW Patient Survey Program, gathering information from patients about their experiences in public hospitals and other healthcare facilities.

BHI publishes a range of reports and tools that provide relevant, accurate and impartial information about how the health system is measuring up in terms of:

- Accessibility – healthcare when and where needed
- Appropriateness – the right healthcare, the right way
- Effectiveness – making a difference for patients
- Efficiency – value for money
- Equity – health for all, healthcare that's fair
- Sustainability – caring for the future.

BHI's work relies on the efforts of a wide range of healthcare, data and policy experts. All of our assessment efforts leverage the work of hospital coders, analysts, technicians and healthcare providers who gather, codify and report data. Our public reporting of performance information is enabled and enhanced by the infrastructure, expertise and stewardship provided by colleagues from NSW Health and its pillar organisations.

bhi.nsw.gov.au