BMJ Open Quality

Continuous individual feedback to nurses at emergency medical dispatch centres: a stepped-wedge, interrupted time series analysis

Douglas Nils Spangler , Hans Blomberg

To cite: Spangler DN, Blomberg H. Continuous individual feedback to nurses at emergency medical dispatch centres: a stepped-wedge, interrupted time series analysis. *BMJ Open Quality* 2025;**14**:e002993. doi:10.1136/bmjoq-2024-002993

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjoq-2024-002993).

Received 5 July 2024
Accepted 11 January 2025

Uppsala Center for Prehospital Research, Institute of Surgical Sciences, Uppsala University, Uppsala, Sweden

Correspondence to
Douglas Nils Spangler;
dnspangler@gmail.com

ABSTRACT

Background Clinical feedback is often lacking in prehospital care, and while performance data is increasingly available to clinical and operational leadership, it is seldom made available to care providers themselves. In this study, we investigate the impact of a simple intervention consisting of the provision of monthly feedback reports via email to emergency medical dispatch nurses in three Swedish regions.

Method Individualised reports consisting of 14 measures divided into descriptive (eg, priority-setting and call times), process (eg, dispatch times and documentation completeness) and outcome (eg, over/under triage rate) categories were developed with staff and management input. Report delivery was implemented using a stepped-wedge design, and effects were evaluated using a hierarchical regression-based interrupted time series analysis.

Results 40 dispatchers were included in the study between March 2020 and October 2023, who handled a total of 246 353 incidents. Some impacts on documentation-related process measures were identified, with the odds of complete documentation increasing by 7.5% (95% CI 5.1 to 9.9) and the odds of having a documented contact reason increasing by 3.8% (1.5–5.9). These effects remained robust over the post-intervention period. Weaker impacts on outcome measures were identified which could be explained by a higher priority given to emergency medical dispatches overall.

Conclusion Providing performance data can influence care providers to adjust their behaviour to improve process-related quality metrics under their direct control. The intervention may also have induced nurses to more often upgrade the priority of their patients. Improving outcome metrics may however require more intensive, multifaceted interventions.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Performance feedback interventions have been found to improve process quality, and to a lesser extent outcomes in prehospital care, but little research has been done in the context of emergency medical dispatching.

WHAT THIS STUDY ADDS

⇒ We evaluate the impact of a continuous, individual-level and automated performance feedback intervention on 14 descriptive, process and outcome measures, finding the strongest impacts on documentation completeness.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study defines and evaluates a suite of quality indicators intended to holistically evaluate quality at an emergency medical dispatching (EMD) centre. We provide the R computer source code and data necessary to not only reproduce our results but also to replicate the intervention at other EMD centres.

INTRODUCTION

Background

Prehospital care is noted for its lack of direct feedback to care providers regarding their performance and patient outcomes. When feedback is provided, it is typically responsive (eg, via incident reporting systems or reviews of specific high-acuity patient types) rather than being systematic in nature.[1 2] The literature on feedback within the field of emergency medical dispatching (EMD) is particularly scant, with researchers in the field of prehospital medicine focusing primarily on ambulance services. While there have been efforts to implement feedback systems in EMD,[3 4] the impact of such systems on dispatcher behaviour is rarely evaluated.[5]

The measurement of healthcare quality has a long history, with Donabedian's structure–process–outcome model representing the most prominent framework for categorising and understanding how aspects of quality interact.[6] In this framework, structural factors account for the conditions which enable care providers to perform processes which contribute to patient outcomes. While structural factors are typically relevant at the system level and thus beyond the control of individual care providers, both process and outcome measures are commonly employed in feedback interventions. Measures of process quality are both the most

commonplace in prehospital care and have been found to be the most susceptible to influence by feedback interventions.[1 7 8] This framework was used as the basis for the selected measures.

Quality measures typically have a normative interpretation such that higher or lower values are considered to represent better quality. Many clinically relevant measures of prehospital care however do not have this type of natural directionality. There is for instance no consensus on the optimal proportion of priority 1 (lights and sirens) dispatches. Such descriptive measures however arguably capture a form of equity—that the likelihood of receiving a priority 1 ambulance response for instance should not depend on which individual call taker handles their call.

There are a number of issues concerning the impact of feedback interventions which must be considered in designing an evaluation. While the identified effects of feedback interventions may be found to be initially strong, effects may diminish over time. Therefore, intervention effects must be measured over a sustained period of time to capture the effects of routinisation.[9] There is furthermore a risk that a focus on certain measures could degrade quality with respect to other unmeasured aspects of quality: measures of under-triage can for instance be easily improved by over-triaging patients without necessarily improving accuracy.[10] Sets of quality measures must therefore be carefully balanced to reduce the potential for care providers to focus on improving the quality of their care as operationalised rather than the underlying quality which the operationalisation is intended to measure.[11] Finally, quality feedback interventions are best implemented with a focus on their application as a learning tool rather than a control mechanism in order to avoid resistance from care providers.[12]

Aim

The aim of this study was to determine how the availability of individual-level feedback on quality indicators to EMD nurses impacted their behaviour in terms of those indicators. Specifically, we hypothesised that the intervention would result in an immediate improvement in the process and outcome measures, and that this improvement would be sustained over time.

METHODS

Study design

An interrupted time series (ITS) analysis of an intervention implemented using a pragmatic stepped-wedge design at three Swedish EMD centres.

Setting

The Swedish regions of Uppsala, Sörmland and Västmanland had a combined population of 983 961 in 2022 and handled about 150 000 emergency calls per year. These regions employ nurses as primary call takers who prioritise all incoming calls to the EMD centre with the aid of a common Clinical Decision Support System (CDSS). Nurses are permitted to deviate from the CDSS recommended priority at their discretion.[13] Nurses may choose to assign an ambulance to a caller with one of four priority levels—1A/1B (lights and sirens response with a higher/lower level of acuity) or 2A/2B (no lights and sirens)—or to not assign an ambulance and instead refer the patient to another healthcare service.[14] The regions employ a common IT system for dispatching from which all data used in this study are drawn.

Prior to this study, the regions had two primary mechanisms in place for quality assurance: incident reporting systems, and a system for monitoring quality measures at the aggregate level. The incident reporting systems are based on reports written by the care providers themselves, other care providers, or patients.[14] While aggregated data had been used to monitor overall system dynamics at the region level, data at the individual level had not been available to care providers about their own performance nor used in evaluating individual-level performance.

Intervention

An approach was sought to furnish care providers with quantitative data about their own performance and triage decisions on a regular basis in a sustainable, low-cost manner. The intervention was developed and implemented by the authors, consisting of a clinically active project manager and the medical director at one of the study sites.

Using the R programming language, a script was written to generate personalised reports based on dispatch system data and e-mail them to nurses once per month.[15] These included a preamble describing rates of missing data affecting the measures, a stacked bar chart describing their call volume by call type and priority level, a table of individual deidentified incidents, and 14 figures detailing the quality measures themselves (see sample report included in the data repository associated with this paper).[16] Each measure was reported for the current month, the average level over the past year, and graphically for 1 year of historical incidents per figure 1 below. Reports were not delivered to nurses who had handled fewer than 50 incidents with care need in a given month to avoid random variation having an undue effect on the reported measures.

The quality measures included in the reports were selected based on previously tracked measures in the regional quality assurance monitoring system, a review of the literature, and consultations with nurses via presentations at staff meetings.[17–20] We sought to include both process measures over which the nurse has a high degree of control (eg, documentation completeness), as well as patient outcome-oriented measures of triage accuracy which can only be controlled indirectly by performing high-quality patient assessments (eg, rates of over-triage and under-triage). While the literature suggests that the former is more likely to result in effective feedback interventions, they are also less directly tied to the core goal of performing rapid and accurate patient assessments.[21] Nonetheless, all measures were developed so as to capture

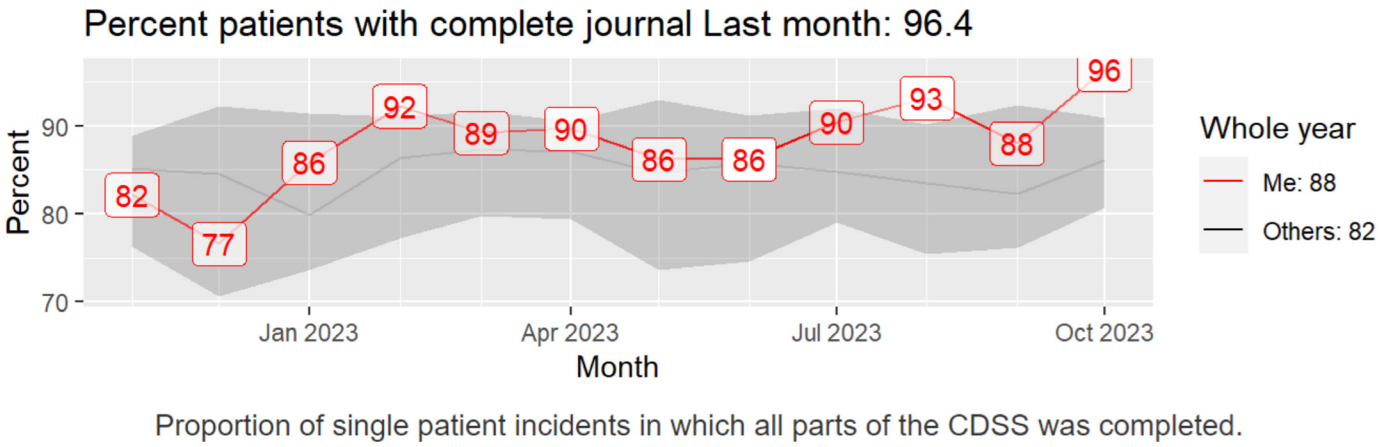

Figure 1 Example measure from report. Red line and values represent the individual nurse's result for the measure (here the proportion of incidents with complete documentation). The grey line and shaded area represent the median and IQR of all other nurses in their region for comparison. The opacity of the label backgrounds corresponds to the number of incidents included in the measure denominator.

aspects of the incident at least indirectly controlled by the nurse (eg, the dispatch time measures relate to only the time taken to identify the need for an ambulance response rather than the total time taken until an ambulance was dispatched), and all measures related only to patients personally assessed by the nurse in question. The development of the measures was however limited by the necessity to draw all data from the dispatching system, as data linkage to ambulance and hospital records was not possible at all study sites.

Prior to implementation in each region, the interpretation of the measures was discussed at staff meetings, emphasising that the feedback reports were intended to be used as a self-improvement tool to enable nurses to see their results in the context of their colleagues. Specifically, it was emphasised that performance regarding descriptive measures such as having a low rate of referral to other care providers compared with the average should not be considered as an inducement to refer more patients, as this could be a patient safety risk. In line with best practices, a monthly report distribution frequency was selected to provide a balance between timeliness of information and to contain a sufficient number of incidents in order to limit the impact of random variation on the interpretation of the measures. Reports were sent from the personal email address of the author, with whom many staff were familiar and had discussed the reports with. The measures in the reports were further crafted with a clear comparator and presented both visually and in text.[21]

The start of the intervention in each region was timed to coincide with the yearly employee interviews in each region so as to allow the reports to be reviewed individually with each nurse during their yearly employee interview to answer any questions. As such, the order of implementation was not randomised, and the intervention was implemented in March 2021 at site A, March 2022 at site B and October 2022 at site C. Beyond their use at yearly employee interviews, there was no mechanism to require nurses to review the reports. The included measures are described in table 1 below.

Statistical analysis

The implementation of the feedback intervention was staggered across the included regions for pragmatic reasons, as well as to allow for analysis following a stepped-wedge design. This design has the benefit of reducing confounding by external events occurring simultaneously with the intervention on the estimation of the intervention effect. An interrupted time series analysis was employed to account for the presence of linear secular trends in the data which would bias a simpler pre-post analysis, and to allow for the estimation of trend changes post-intervention to address the research questions related to the effect of the intervention over time.[22] All nurses who handled at least 50 incidents per month during at least six of the 12 months pre-intervention and post-intervention in their region were included in the analysis. In accordance with the ethics approval for this study, data from nurses who opted out of participation in the study were excluded from the research dataset. Call data from included nurses was collected from 1 year prior to the intervention at the first site to 1 year after intervention at the last site. The 1 year pre-intervention time frame was selected so as to avoid including pre-COVID-19 data in the estimate of secular trends which, based on inspection of longer-term internal quality assurance data, would violate the assumption of linearity. In this analysis, coefficients were included to estimate pre-intervention (secular) trends, immediate post-intervention level changes, and a post-intervention trend to capture the longer-term impacts of the intervention. Given that we could not verify that nurses read the delivered reports, this study could be considered as targeting an 'intention-to-treat' estimand rather than the effect of the information contained in the reports per se.

A hierarchal regression approach was used to estimate effects using individual nurses as the unit of analysis using

Table 1 Description of measures included in feedback reports

Measure group	Measure name	Description
Descriptive	(D1) Percent priority 1	Proportion of all incidents involving a medical problem receiving a priority 1 (lights and sirens) response.
	(D2) Percent referred	Proportion of all incidents involving a medical problem who did not receive an ambulance response.
	(D3) Percent manually downgraded patients	Proportion of all single patient incidents (excluding for example, planned transfers and multipatient incidents) where the response priority was lower than that suggested by the CDSS.
	(D4) Percent manually upgraded patients	Proportion of all single patient incidents where the response priority was higher than that suggested by the CDSS.
	(D5) Median call time	Median time from answering to ending the first call associated with all incidents.
Process	(P1) Percent patients with life threat questions	Proportion of single patient incidents with initial questions intended to exclude life-threatening conditions completed (breathing/airway clear/conscious).
	(P2) Percent patients with contact reason	Proportion of single patient incidents with a documented reason for contact in the CDSS.
	(P3) Percent patients with complete journal	Proportion of single patient incidents in which all parts of the CDSS were completed.
	(P4) Median queue time	Median time from receipt of call at the EMD centre to answering the first call associated with all incidents.
	(P5) Median dispatch time priority 1A	Median time from answering the call to the first of: marking the call as ready for dispatch, connecting a dispatcher to the call to handle ambulance dispatch, actual ambulance dispatch for priority 1A (immediately life threatening) incidents.
	(P6) Median dispatch time priority 1B	Same as above for priority 1B (emergency, non-life threatening) incidents.
Outcome	(O1) Percent priority 1 patients not transported	Proportion of incidents dispatched as priority 1A or B which were treated on scene and not transported to a hospital (excluding call types considered to be appropriately treatable on scene: cardiac arrest, allergic reaction, low blood sugar, breathing difficulty in children 5 or under). A measure of over-triage.
	(O2) Percent priority 2 patients transported priority 1	Proportion of incidents dispatched as priority 2A or B which were subsequently transported by the ambulance on scene to a hospital with lights and sirens. A measure of under-triage.
	(O3) Percent referred patients transported within 72 hours	Proportion of single patient incidents referred to non-ambulance care where the patient subsequently re-contacted the EMD centre within 72 hours and was transported to the hospital. A measure of under-triage.

CDSS, Clinical Decision Support System; EMD, emergency medical dispatching.

random intercepts at the individual nurse level to account for repeated measures.[23] Binomial models were used to estimate effects regarding proportion-based measures, and weighted linear models with log-transformed dependent variables were used for skewed time-interval-based measures. Sensitivity analyses evaluating a dataset excluding the first wave of the COVID-19 pandemic and excluding the month immediately prior to the intervention to remove anticipatory effects were performed and are reported in online supplemental material 1. Customary 95% CIs for parameter estimates are reported based on 1000 parametric bootstraps, but results are interpreted holistically, taking into account sensitivity analyses and the context of the multiple comparisons performed in this study. The possibility of homogenising effects (ie, nurses regressing to the mean of their colleagues) was not assessed, although a visual inspection of measures with regard to their variability pre-intervention and post-intervention at the individual nurse level was performed. Scatterplots displaying individual nurse performance pre-intervention and post-intervention did not demonstrate clear reductions in variability (see online supplemental material 1), but the presence of regression to the mean effects cannot be excluded.

All analyses were performed using R (v 4.3.1) and the lme4 package to fit models.[24 25] The code for all analyses performed in this article may be reviewed in the analysis notebook included as online supplemental material 1

Table 2 Descriptive data

	Site A	Site B	Site C	Total
Number of included nurses	13	15	12	40
Median nurse age	43	46	47	45
Percent female nurses	78.2	61.5	60.6	65.2
Number of included incidents	86 453	78 690	81 210	246 353
Median patient age	66.7	69.5	71.5	69.4
Percent female patients	52.1	50.8	51.0	51.3

and may be reproduced using the R source code and data included in the associated data repository.[16] All scripts necessary to replicate the intervention are available on Github.[15] The study was approved by the Swedish Ethics Review Authority (dnr 2021-06340).

RESULTS

A total of 651 091 incidents of all types were identified in the dispatch system database from 1 year prior to intervention in the first site (2020-03-01) to 1 year after intervention in the last site (2023-10-01). Of these, 457 262 were incidents involving a medical problem, with the remainder consisting of duplicates, police/fire referrals, stand-bys, etc. These incidents were handled by a total of 210 staff, of whom 42 had handled 50 or more incidents per month during six of the 12 months prior to and after the intervention in their region. The remainder consisted of nurses employed only prior to or after the intervention, part-time staff, and dispatchers only handling non-patient incidents. Two nurses were excluded due to not consenting to participation in the study. A total of 40 nurses handling 246 353 incidents were thus included, distributed per table 2. Incidents were relatively evenly distributed by site, although site B had fewer incidents per included nurse. Median patient age was 69.4 with slightly more female than male patients.

Intervention effects for each of the 14 metrics are presented visually in figure 2 as median values and IQRs. Quantitative analysis results from the ITS regression analysis are presented in table 3. Estimates are presented as one minus the odds ratio for binomial models, and multiplicative effects (back-transformed logarithmic proportions) for log-normal (time) models, both of which are interpreted with a value of 0 as a null effect.

Regarding the descriptive measures, some effects may be discerned despite not explicitly defining a desired effect direction. The odds of manually upgrading a patient (D4) showed an increase of 4.8% which was statistically significant at the 95% confidence level, while all other measures regarding priority-setting reflect a trend towards higher priority levels post-intervention. Notably, these effects occurred in the context of a pre-intervention trend towards lower priority-setting.

The process measures showed the most substantial intervention effects, with the odds of a patient call having a documented contact reason (P2) increasing by 3.8% and the odds of a patient having a complete journal (P3) increasing by 7.5%. The related measure of documented life threat questions (P1) showed no significant changes. The average post-intervention trend for these measures was positive, suggesting a sustained impact of the feedback reports over time. These effects were not uniform across all study sites however, and site A had a substantially lower pre-intervention level and larger effect size regarding documentation completeness, and an apparent decreasing trend in journal completeness during the post-intervention time frame (see online supplemental file 1 for per-region analyses). We found no significant effects regarding the time-based process measures (P4–P6). We also found no significant effects among the outcome measures, although both over-triage and under-triage measures reflected the overall trend towards higher priority-setting, with trends towards both higher levels of over-triage (O1) and lower levels of under-triage (O2/O3).

A number of sensitivity analyses were conducted to evaluate the robustness of the findings in the main analysis. These included an estimation of separate models for each site as noted previously to evaluate between-region heterogeneity of effects. Given that the design of the feedback reports was presented to staff during the month prior to their distribution, it is possible that there are some spillover effects prior to intervention, and signs of such anticipatory effects may be seen in figure 2. To evaluate the impact of this, models were estimated excluding data from the months prior and during the intervention. The results indicate robustness regarding the findings relating to documentation completeness, and similar results regarding patient outcomes. The pre-intervention time frame includes the COVID-19 pandemic, the first wave of which had a substantial impact on the triage process in the studied regions. To evaluate the impact of this, the first 3 months of the dataset were excluded in a sensitivity analysis suggesting results similar to those found in the main analysis. Visual inspection of scatterplots displaying individual nurse performance pre-intervention and post-intervention did not reveal obvious reductions in variability (see online supplemental material 1), but the presence of more subtle effects potentially identifiable by quantitative analysis cannot be ruled out.

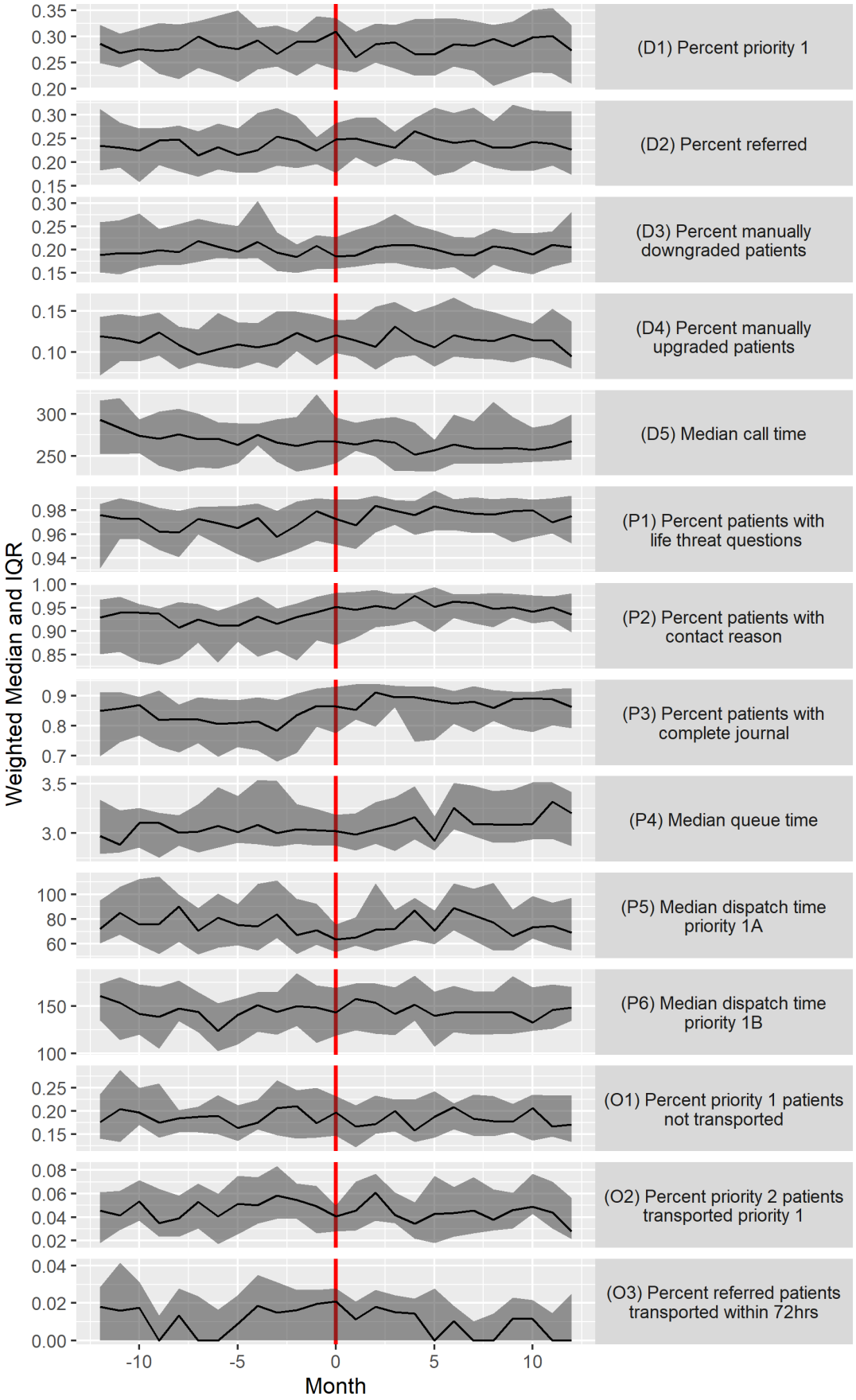

Figure 2 Time series plots. Median and IQR values for all included nurses (weighted by the number of incidents constituting the denominator of each measure) are plotted for 1 year pre-intervention and post-intervention for all study sites, aligned such that the intervention occurs at month 0. All time measures are expressed in seconds.

Table 3 Interrupted time series coefficient table

Measure group	Measure name	Pre-intervention trend (per year)	Level change	Post-intervention trend (per year)
Descriptive	(D1) Percent priority 1	**–2.2 (–4.1:–0.2)**	2.4 (–0.5:5.5)	–0.5 (–3.2:2.2)
	(D2) Percent referred	**7.4 (4.9:9.6)**	–1.5 (–4.4:1.8)	–1.7 (–4.5:1.0)
	(D3) Percent manually downgraded patients	**5.5 (3.2:8.1)**	–3.2 (–6.5:0.5)	**–4.1 (–7.2:–1.0)**
	(D4) Percent manually upgraded patients	**–6.6 (–9.3:–3.7)**	**4.9 (0.4:9.7)**	**8.8 (4.2:13.3)**
	(D5) Median call time	–2.1 (–11.5:9.9)	1.6 (–17.2:25.1)	–0.8 (–15.2:16.5)
Process	(P1) Percent patients with life threat questions	–0.1 (–1.3:1.1)	1.1 (–1:3.2)	–0.0 (–1.7:1.6)
	(P2) Percent patients with contact reason	**–2.1 (–3.4:–0.8)**	**3.8 (1.6:6.1)**	**2.6 (0.77:4.5)**
	(P3) Percent patients with complete journal	**–5.7 (–7.1:–4.3)**	**7.5 (5.0:10.1)**	**5.9 (3.9:8.2)**
	(P4) Median queue time	2.3 (–5.8:11.1)	0.7 (–13.4:16.0)	–1.0 (–12.3:11.2)
	(P5) Median dispatch time priority 1A	3.2 (–5.3:12.2)	–3.0 (–16.6:12.1)	–0.5 (–11.8:12.4)
	(P6) Median dispatch time priority 1B	–0.6 (–9.9:11.4)	2.8 (–15.9:23.2)	–0.1 (–13.94:15.7)
Outcome	(O1) Percent priority 1 patients not transported	**–11.4 (–14.4:–8.1)**	4.5 (–1.6:11.3)	**11.0 (5.4:16.8)**
	(O2) Percent priority 2 patients transported priority 1	**9.5 (2.3:16.2)**	–10.1 (–19.3:0.2)	1.3 (–7.2:11.4)
	(O3) Percent referred patients transported within 72 hours	–5.2 (–13.3:2.9)	–10.5 (–23.6:2.6)	–5.9 (–16.9:6.1)

Effect sizes are expressed as percentage changes in ORs for binomial outcomes, and as multiplicative percentage changes for continuous outcomes (modelled as log-linear). Estimates whose 95% CI exclude a null effect are bolded.

DISCUSSION

In this study, we identified modest but robust effects of a simple continuous feedback intervention on process measures relating to documentation completeness. Post-intervention trends furthermore suggested that the effects that were found were sustained over time, with average post-intervention trends suggesting an improvement over time regarding documentation completeness. However, we found little impact on outcome measures, or process measures relating to queue or dispatch times. The trends in the outcome measures could furthermore be explained by changes in overall priority-setting rather than any improvement in triage ability. Nurses tended to assign higher priorities to patients in the post-intervention period as reflected by overall trends in the descriptive and outcome measures, which could be a reflection of an increased focus on patient safety over resource utilisation efficiency on receiving the reports. These results highlight that quality measures are often interdependent, and must be considered as a whole, rather than focusing on specific measures.

These results are in line with previous research suggesting that feedback interventions in the context of ambulance services have the greatest potential to improve process measures such as documentation and protocol adherence, and weaker effects regarding impacts on patient outcomes.[1 2] There are notable exceptions to this, even among studies evaluating simple 'informational' interventions,[26] although substantial positive impacts on patient outcomes are more typically found for comprehensive interventions combining feedback and education for specific high-acuity conditions such as cardiac arrest.[27 28] A performance feedback intervention in the EMD setting was found to improve protocol adherence,[5] although this study took place over a relatively short time frame, and did not investigate subsequent patient outcomes.

Whereas many quality improvement initiatives are complex, multifaceted projects which require substantial resources to implement and sustain, one goal of the intervention investigated in this study was to keep it simple and low-cost. After the initial work to develop the feedback system, regularly delivering the reports and performing basic data quality checks requires about 2–3 hours of work per month and has no other associated costs. As such, even the modest effects on processual quality measures found here may therefore suffice to consider the intervention cost-effective. The intervention is technically simple for others to implement given the availability of appropriate data, and all source code necessary to replicate the intervention is available on GitHub.[15]

To extend the effects, additional interventions could be tied to this form of feedback to improve the ability to affect not only highly controllable process measures but also outcomes related to triage accuracy. Performance with regard to outcome measures within specific patient categories (eg, trauma, children, stroke, etc) could for instance be used to select staff for participation in targeted continuing medical education efforts. Another potential improvement could be to select rare but significant indicators of potential patient harm (eg, patients referred to non-emergency care with a high priority ambulance response within 24 hours) for mandatory review by either

the nurse themselves or clinical leadership rather than relying on voluntary review of the reports. Such additional interventions have the potential to improve outcome measures by increasing the timeliness and specificity of the feedback information, and providing a clearer path to addressing deficiencies which are often missed by traditional incident reporting systems, particularly with regard to patient outcomes.[14]

Limitations

The main threat to the validity of interrupted time series analyses such as this is history bias, that is, events occurring simultaneously with the intervention and confounding the results. This risk is limited in the study design by the staggered implementation of report delivery across the sites. Another consideration in interpreting these results is the potential effects of observation bias, that is, the effect of knowing that one's performance is being observed. In the context of feedback interventions however, the observer effect could be argued as being part of the intervention effect, as the reports are in a sense a tool intended to communicate the priorities of an organisation to its staff. Such observer effects are likely to be most prominent at study site A, which is where the research group is based, and where the intervention effects indeed were the strongest. These varying contextual effects may impact the generalisability of the results, and also emphasise the beneficial effect of implementers working closely with clinicians. A further limitation is that the implementation order was not randomised but rather for pragmatic reasons followed the order in which the regions performed their yearly employee interviews.

The ITS analysis involves an assumption of linear trends over the pre-intervention and post-intervention time frames. This assumption appears to hold based on visual inspection, and results were not impacted by excluding the period containing the most prominent threat to this assumption (the first wave of the COVID-19 pandemic). While the study included only nurses who regularly received feedback reports via email, we lack data regarding how often the included nurses actually reviewed the feedback reports, which could be investigated in the future for instance using surveys or by obtaining read receipts from staff. Their use as the basis for yearly employee interviews entails that all included nurses were at least aware of the availability of the reports. A final limitation regards the included performance measures themselves. Ideally, patient outcomes would include more objective measures such as mortality and care provided by other parts of the healthcare system, but only data from within the ambulance dispatching system itself was available in all studied regions.

CONCLUSION

This relatively simple feedback intervention showed modest but robust effects regarding quality measures over which nurses had the most control, namely measures of documentation completeness. Measures related to patient outcomes over which nurses had less control did not see improvements. More complex and intensive interventions tied to feedback may be necessary to impact patient outcomes.

Contributors DNS and HB jointly conceived of and designed the intervention and study. DNS implemented the intervention, performed the analysis and drafted the manuscript. DNS and HB jointly reviewed and revised the manuscript. HB is the guarantor of the study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data and scripts necessary to reproduce the results as reported here are available on the Open Science Foundation platform as referenced in the manuscript.

ORCID iD

Douglas Nils Spangler http://orcid.org/0000-0001-6775-5051

REFERENCES

1 Eaton-Williams P, Mold F, Magnusson C. Effective clinical feedback provision to ambulance clinicians: a literature review. *J Paramed Pract* 2020;12:109–17.
2 Wilson C, Janes G, Lawton R, *et al*. Types and effects of feedback for emergency ambulance staff: a systematic mixed studies review and meta-analysis. *BMJ Qual Saf* 2023;32:573–88.
3 Lindström V, Falk AC, Castrén M. Feedback on emergency medical dispatching – the ambulance crew experience of a computer based feedback system. *Scand J Trauma Resusc Emerg Med* 2009;17:O4.
4 Klesse T, Krog J, Mikkelsen S. Accuracy of judgment of urgency as perceived by emergency medical dispatch centres in Denmark- a quality assurance study. *Resuscitation* 2019;142:e75.
5 Clawson JJ, Cady GA, Martin RL, *et al*. Effect of a comprehensive quality management process on compliance with protocol in an emergency medical dispatch center. *Ann Emerg Med* 1998;32:578–84.
6 Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;260:1743–8.
7 Ivers N, Jamtvedt G, Flottorp S, *et al*. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;2012:CD000259.
8 Knaup C, Koesters M, Schoefer D, *et al*. Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *Br J Psychiatry* 2009;195:15–22.
9 Simon M, McGuire R, Lynch H, *et al*. The Journey Toward Routinization: Triage Nursing and the Success of an Emergency Department-Based Routine HIV Testing Program. *J Emerg Nurs* 2016;42:183–5.
10 Lupton JR, Davis-O'Reilly C, Jungbauer RM, *et al*. Under-Triage and Over-Triage Using the Field Triage Guidelines for Injured Patients: A Systematic Review. *Prehosp Emerg Care* 2023;27:38–45.
11 Lester HE, Hannon KL, Campbell SM. Identifying unintended consequences of quality indicators: a qualitative study. *BMJ Qual Saf* 2011;20:1057–61.
12 Báo ACP, Amestoy SC, Moura GMSS de, *et al*. Quality indicators: tools for the management of best practices in Health. *Rev Bras Enferm* 2019;72:360–6.
13 Holmström IK, Kaminsky E, Lindberg Y, *et al*. Registered Nurses' experiences of using a clinical decision support system for triage

of emergency calls: A qualitative interview study. *J Adv Nurs* 2020;76:3104–12.
14 Spangler D, Edmark L, Winblad U, *et al*. Using trigger tools to identify triage errors by ambulance dispatch nurses in Sweden: an observational study. *BMJ Open* 2020;10:e035004.
15 Spangler D. Github repository: dnspangler/feedback_reports. 2024. Available: https://github.com/dnspangler/feedback_reports [Accessed 31 Jan 2024].
16 Anonymous. Continuous feedback dataset. 2024. Available: https://osf.io/g8kc3/ [Accessed 14 Feb 2024].
17 Khorram-Manesh A, Lennquist Montán K, Hedelin A, *et al*. Prehospital triage, discrepancy in priority-setting between emergency medical dispatch centre and ambulance crews. *Eur J Trauma Emerg Surg* 2011;37:73–8.
18 Turner J, Siriwardena AN, Coster J, *et al*. Developing new ways of measuring the quality and impact of ambulance service care: the PhOEBE mixed-methods research programme. *Programme Grants Appl Res* 2019;7:1–90.
19 Clawson J, Olola CHO, Heward A, *et al*. Accuracy of emergency medical dispatchers' subjective ability to identify when higher dispatch levels are warranted over a Medical Priority Dispatch System automated protocol's recommended coding based on paramedic outcome data. *Emerg Med J* 2007;24:560–3.
20 Jensen JL, Carter AJE, Rose J, *et al*. Alternatives to Traditional EMS Dispatch and Transport: A Scoping Review of Reported Outcomes. *CJEM* 2015;17:532–50.
21 Brehaut JC, Colquhoun HL, Eva KW, *et al*. Practice Feedback Interventions: 15 Suggestions for Optimizing Effectiveness. *Ann Intern Med* 2016;164:435–41.
22 Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol* 2017;46:348–55.
23 Fok CCT, Henry D, Allen J. Research Designs for Intervention Research with Small Samples II: Stepped Wedge and Interrupted Time-Series Designs. *Prev Sci* 2015;16:967–77.
24 Bates D, Mächler M, Bolker B, *et al*. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* 2015;67:1–48.
25 R Core Team. R: a language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2019. Available: https://www.R-project.org/
26 Persse DE, Key CB, Baldwin JB. The effect of a quality improvement feedback loop on paramedic-initiated nontransport of elderly patients. *Prehosp Emerg Care* 2002;6:31–5.
27 Tanaka Y, Taniguchi J, Wato Y, *et al*. The continuous quality improvement project for telephone-assisted instruction of cardiopulmonary resuscitation increased the incidence of bystander CPR and improved the outcomes of out-of-hospital cardiac arrests. *Resuscitation* 2012;83:1235–41.
28 Hopkins CL, Burk C, Moser S, *et al*. Implementation of Pit Crew Approach and Cardiopulmonary Resuscitation Metrics for Out-of-Hospital Cardiac Arrest Improves Patient Survival and Neurological Outcome. *J Am Heart Assoc* 2016;5:e002892.