Multi-method appraisal of clinical quality indicators for the Emergency Medical Services in the Low- and Middle-Income Setting: The South African Perspective

Ian Howard (ianhoward@outlook.com)
Karolinska Institutet
Peter Cameron
Monash University
Maaret Castrén
Helsinki University
Lee Wallis
University of Cape Town
Veronica Lindström
Karolinska Institutet

Research article

Keywords: multi-method appraisal, clinical quality, quality indicator

Posted Date: January 14th, 2020

DOI: https://doi.org/10.21203/rs.2.20860/v1

Abstract

Background: Quality Indicator (QI) appraisal protocols are a novel methodology that combine multiple appraisal methods in order to comprehensively assess the "appropriateness" of QIs for a particular healthcare setting. However, they remain inadequately explored compared to the single appraisal method approach. The aim of this paper was to describe and test a QI appraisal protocol versus the single method approach, against a series of QIs potentially relevant to the South African Prehospital Emergency Care setting.

Methods: An appraisal protocol was developed consisting of two categorical-based appraisal methods, the Qualify tool and Rand/Appropriateness method, combined with the qualitative analysis of the discussion generated during the consensus application of each method, by a QI Appraisal Working Group. Inter-rater reliability of each individual method was assessed prior to group consensus rating. Variation in the number of non-valid QIs identified between each method and the proportion of non-valid QIs identified between each method and the protocol were compared and assessed.

Results: There was mixed inter-rater reliability of the individual methods prior to the group consensus. There was similarly poor to moderate correlation of the results obtained between the individual methods (Spearman's rank correlation = 0.42, $p<0.001$). From a series of 104 QIs, 11 were identified that were shared between the appraisal methods. A further 19 QIs were identified and not shared by each method, highlighting the benefits of a multimethod approach. There was little evidence to support a difference in the proportion of non-valid QIs identified between individual methods (difference=-0.03); between the Quality tool and the protocol (difference=-0.05); or between the Rand method and the protocol (difference=-0.02). The outcomes were additionally evident in the group discussion analysis, which in and of itself added further input towards understanding and appraising the appropriateness of the QIs that would not have otherwise been captured or understood by the individual methods alone.

Conclusion: The utilisation of a multi-method appraisal protocol offers multiple benefits, when compared to the single appraisal approach, and can provide the confidence that the outcomes of the appraisal will ensure a strong foundation on which the measurement framework can be QI successfully implemented and employed.

Background

The Institute of Medicine defines healthcare quality as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge"[1]. Objectively assessing the extent to which this is achieved can be a challenging task, given that quality is a relatively abstract concept. It stands to reason therefore that a central tenet to defining quality is the system or framework used towards its measurement.

The measurement of healthcare quality provides an essential mechanism towards informing policy and directing strategy; identifying and benchmarking performance; guiding priorities[2–4] and improvement initiatives; and maintaining transparency and accountability of the system. Consequently, multiple users will consume quality data in a variety of ways in order to achieve these aims. Therefore, for any system of quality measurement to be successful, it is fundamental that it be comprehensive in its approach, yet simple in its design, and contextually relevant in order to provide an appropriate measure of quality.

Considerable progress has been made towards improving Prehospital Emergency Care (PEC) quality measurement, largely in the form of the development of PEC-specific quality indicators (QIs)[5–7]. In and of themselves, QIs cannot improve quality, they effectively act as measurement tools, screens, or flags that provide clinicians and organisations with a quantitative basis to monitor, evaluate, and improve the quality of patient care, clinical support services, and organizational function[3,4]. Despite their advantages, the objective appraisal of such systems of quality measurement is often neglected, leading to the potential for implementation of inappropriate QIs. In the PEC environment, this is already evident in the literature where less than 15% of PEC focused QIs have undergone some form of comprehensive measure evaluation[7]. The consequences of inadequately assessed reliability, validity and bias in quality measurement can in the best-case scenario prove to be time-consuming and costly, and in worst case scenario potentially undermine the system in its entirety and impact patient safety[8,9].

Several individual methodologies to appraise QIs have been described and utilised with considerable success[9–17]. While there is a level of overlap or commonality in the components that they assess, the process towards their application can vary significantly[9–17]. Therefore, the potential exists for variation in the outcomes of these methodologies when applied to a common data set. QI appraisal protocols are a novel methodology that combine multiple appraisal methods in order to comprehensively assess the "appropriateness" of QIs for a particular healthcare setting[9]. There is evidence, albeit limited, to suggest benefits with the use of such protocols, however they remain inadequately explored as an option compared to the single appraisal method approach. The aim of this paper was to therefore describe and test a multi-method multi-part QI appraisal protocol versus the single method approach, against a series of QIs previously identified as potentially relevant to the South African (SA) PEC setting.

Methods

The triangulation and integration of multiple data types has been increasingly recognised as a valuable approach towards the study of healthcare delivery[18–20]. For the purposes of this study, an appraisal protocol was developed consisting of two categorical-based appraisal methods, combined with the qualitative analysis of the discussion generated during the consensus application of each method, by a QI Appraisal Working Group (Fig. 1). For part 1, the Qualify QI appraisal tool was used consisting of four-level Likert scale questions (1 = Does not apply; 2 = Rather does not apply; 3 = Rather applies; 4 = Applies) to assess 18 criteria amongst three categories: Relevance (3 criteria); Scientific Soundness (6 criteria) and Feasibility (9 criteria) (Table 1). The appraisal tool employed was selected based on its previous use to assess the appropriateness of QIs and given its focus on feasibility[11,12].

For part 2, the Rand/UCLA Appropriateness Method was used to further rate the indicators by testing the definitions, data components and criteria for use developed for each QI against several clinical vignettes. Four categories (Clarity, Necessity, Acceptability and Technical Feasibility) were rated using a 9-point visual analogue scale, and data extraction assessed using a mock-up of a generic patient report form for the clinical vignettes[9,13]. Two separate vignettes were developed for each of the QI categories included in the data extraction, and a "low quality documentation" and "high quality documentation" version developed for each vignette for use during the assessment. The Rand method has previously been utilised to assess QIs and was additionally included based on its practical focus (i.e.: the data extraction)[14–17].

Both methods consisted of an evidence evaluation component as part of the appraisal process. To achieve this, the QIs were assessed for inclusion within local clinical practice guidelines (CPGs), and against the results of a literature review of the evidence base utilised for the development of PEC focused QIs. The results of the review were assessed and presented using the Oxford Centre for Evidence-based Medicine Levels of Evidence[21].

Data for parts 1 and 2 were collected over three rounds of group discussion of a QI Appraisal Working Group, facilitated by the principle investigator (IH). An initial introductory round was conducted to familiarize the Working Group with the appraisal tool, Rand methodology, results of the literature review, and provide the data dictionary for the QI set. Prior to Round 1, the appraisal tool was independently applied by each member of the Working Group, who then met to discuss their individual scoring and apply a final consensus summary score for Round 1. Prior to Round 2, the Working Group similarly independently assessed the results of the literature review, and then met to apply a final consensus rating of the evidence. Round 2 was further utilised to introduce the clinical vignettes for each category which would be utilised for the data extraction. For round 3, the Working Group met to compare their individual data extraction results and rate the QIs for the categories of the Rand method. The Working Group meetings were recorded and later transcribed for the final part of data collection – content analysis of the discussion generated surrounding the consensus appraisal process for Rounds 1 to 3.

Setting and Population

Traditionally, quality in the PEC setting in SA has been exclusively reported based around response time targets[22 – 25]. Utilisation and reporting of clinically focused QIs by the Emergency Medical Services (EMS) in SA is wholly lacking. Towards this, several clinically focused PEC QIs have recently been identified for potential relevance to the SA PEC setting[26]. As a result, these QIs were used to test the appraisal protocol, with the secondary aim of identifying those QIs appropriate for use in the SA PEC setting.

The QI Appraisal Working Group consisted of nine preselected experts chosen for their intricate knowledge of the South African PEC setting and to align with minimum panel size recommendations for each methodology[10,13]. All the participants were South African trained and post-graduate educated Emergency Care Practitioners (ECPs) with > 10 years operational experience each. Six of the participants' primary experience and occupation were in quality governance and improvement within PEC in general, and the remaining three were primarily involved in clinical operations. The Working Group were given one month between each round with which to work through the information and data collection required for each subsequent round.

Data Analysis

Descriptive statistics were utilised to describe and summarize the categorical based appraisal data. For the appraisal tool, mean scores per category, and the number of criteria scoring either 3 (Rather applies) or 4 (Applies) were calculated and presented for each QI. For the Rand method rating, consensus scores per category, and the proportion of categories scoring 7 or more were calculated and presented for each QI. Despite the face-to-face consensus process, inter-rater reliability for each criterion of both the appraisal tool and Rand method were calculated using percentage agreement and Gwet's AC1 and presented for reporting purposes.

A final composite score was calculated for each QI, for each method. For the appraisal tool, this was calculated using a weighted mean of the appraisal categories after consensus, due the differences in number of criteria per category. To be considered a valid indicator, the QI had to score ≥ 3 based on the final composite score. For the Rand method, the unweighted mean of the appraisal categories after consensus was used. To be considered a valid indicator, the QI had to score ≥ 7 based on the final composite score. A second group of QIs were identified consisting of those scoring on the validity threshold (3.0-3.1 for the Quality tool; 7.0-7.1 for the Rand method) for which caution was recommended prior to full implementation.

Correlation between the final composite scores of each method for each QI was calculated and presented using the Spearman's rank correlation. The consensus derived proportion of non-valid QIs, and QIs for which caution was recommended, identified by each individual method and the protocol, were calculated and assessed against each other using the z test. 95% confidence intervals will be calculated where necessary and a p-value of 0.05 used as a cut-off for strength of evidence. All data were entered and analysed using a combination of Microsoft Excel 2010 (Microsoft Corp., Richmond, WA, USA) and Stata version 16 (StataCorp. College Station, TX: StataCorp LLC).

Conventional content analysis, as described by Hsieh and Shannon, was utilised to sort and analyse the group discussions generated during the three rounds[27]. Recordings and transcripts were created for each round, and each transcript reread for content familiarisation. First-level coding was conducted through the extraction of meaning units from each transcript and summarised into codes using open-coding from each interview. Once completed, similar codes were combined and organised to develop clustered sub-categories pertaining to each appraisal tool. Transcriptions were analysed using MAXQDA software for data storage; extraction of meaning units and sub-category development (MAXQDA, 2016; Sozialforschung GmbH, Berlin, Germany).

Results

The Working Group appraised a total of 90 *clinical* and 14 *non-clinical* (n=104) QIs using each method, over the three rounds. There was a high level of validity of the QIs assessed across the majority of the appraisal criteria for both methods, the results of which were poor to moderately correlated between each

method.

Round 1 - QI Appraisal Tool

There was mixed inter-rater reliability of the criteria found prior to the group consensus discussion. General *Validity* and *Understandability and interpretability for medical and nursing personnel* scored perfect agreement within the group, while *Data Collection Effort* (% agreement = 22%, IRR = 0.01) and *Understandability and interpretability for patients and interested public* (% agreement = 28%, IRR = 0.09) and scored the lowest (Table 2) . Of the 104 QIs assessed, eight (7.7%) scored less than the validity threshold on the final composite score (≥3). Four of these were in the Acute Coronary Syndromes (ACS) clinical category, and six were QIs associated with or influenced by a receiving facility or location. All eight of these QIs scored relatively high for *Relevance* and *Scientific Soundness* yet scored poorly for *Feasibility*. On average, overall scores were generally higher for criteria within *Relevance* and *Scientific Soundness* and lower for those within *Feasibility* (Table 3). Another 15 QIs scored on the validity threshold (3.0 - 3.1) and were generally associated with resources/equipment or regarding the identification and reporting of sentinel events. These QIs similarly had their overall score reduced due to a reduced perception of potential feasibility.

For the purposes of appraising the *Indicator Evidence* criterion within the *Scientific Soundness* category, the QIs were evaluated for inclusion within local clinical practice guidelines (CPGs). There was considerable representation of the QIs amongst the SA national EMS CPGs (Table 3). Seventy-nine QIs (76%) were accounted for in the CPGs, of which 76 (73%) had evidence directly supporting their use. Those QIs not represented were found to be either structure based QIs; clinical bundle based QIs; or those QIs focusing on sentinel events and patient safety.

Round 2 – Literature Review

The literature search identified a total of 1624 potential articles for review (Figure 2). Following the title and abstract review, 1528 articles did not meet inclusion criteria and were excluded, leaving 89 articles for full-text review. An additional 15 articles were included following a review of the list of references of the 96 articles identified. Following the removal of duplicate texts, and research not meeting the inclusion criteria (n=57) 31 articles remained for the full-text review. The literature review found an evidence base for 11 of the 15 Clinical subcategories and the 2 Non-clinical subcategories, plus an additional 4 subcategories not included in the QI appraisal, covering 311 indicators (Table 4). In excess of half (59%) were developed through a consensus/expert opinion-based approach, with fewer developed via more robust and higher quality levels of evidence such as systematic reviews and/or cohort and case control-based studies (10% each).

Round 3 – Rand Method

As with the appraisal tool, there was mixed inter-rater reliability in the individual rating prior to the consensus rating, with *Acceptability* scoring the highest (% agreement = 90%, IRR = 0.9) and *Technical Feasibility* the lowest (% agreement = 47%, IRR = 0.32). Eleven QIs (10.6%) scored below the validity threshold, six of which were within the ACS clinical category, including the four identified using the appraisal tool. Similarly, the same six QIs associated with a receiving facility or location scoring below the validity threshold with the appraisal tool, scored below the validity threshold using the Rand method. Another eight QIs scored on the validity threshold (7.0 - 7.1) and were generally associated with resources/equipment. Only two of these QIs matched those scoring on the threshold with the appraisal tool. Again, as with the appraisal tool, scores were lower within the *Technical Feasibility* category compared to the other three.

Comparison of Categorical Appraisal Methods

When final consensus validity scores were compared, there was poor to moderate correlation of the results obtained between the appraisal tool and Rand method (Spearman's rank correlation = 0.42, p<0.001). Ninety-two of the 104 QIs (88%) (78 *clinical* and 14 *non-clinical*) were appraised to be valid and feasible for the SA PEC setting, based on the results of this study. Of this group, an additional 21 QIs (13 *clinical* and eight *non-clinical*) were assessed to be on the threshold of validity, in which caution is recommended until a pilot study on their use can be conducted, prior to full implementation. There was little evidence to support a statistical difference in the proportion of non-valid QIs identified between the Qualify tool and the Rand method [difference = -0.03; (95% CI -0.12:0.05, p=0.47)]; between the Quality tool and the protocol [difference = -0.05; (95% CI -0.13:0.03, p=0.25)]; or between the Rand method and the protocol [difference = -0.02; (95% CI -0.11:0.07, p=0.66)]. There was likewise little evidence to support a statistical difference in the proportion of QIs in which caution is recommended, identified between the Qualify tool and the Rand method [difference = 0.07; (95% CI -0.02:0.15, p=0.12)]; or between the Quality tool and the protocol [difference = -0.06; (95% CI -0.16:0.04, p=0.27)]. There was however, strong evidence to support a statistical difference between the proportion of QIs in which caution is recommended, identified between the Rand method and the protocol [difference = -0.13; (95% CI -0.22:-0.03, p=0.009)].

Discussion Group Content Analysis

Several observations highlighted during the group discussions were found to be important considerations regarding the appraisal protocol and its ability to assess the appropriateness of the QIs for the SA PEC setting. For the appraisal tool, *Relevance* and *Scientific Soundness* were perceived to be characteristics inherent to the QIs (and supporting data components) themselves, and as a result were generally appraised to be highly applicable across all QIs and criteria (Table 5). In contrast, *Feasibility* was judged to be more of a gauge of the system in which the QIs would be implemented and as such, scores were found to be on average lower amongst these criteria [1.1, 1.2]. Somewhat related to this, was the broader issue of context and the importance of selecting those indicators that best suited the local setting, prior to full implementation (1.3, 1.4]. Despite the focus on the appraisal of the QIs, on several occasions the discussion steered towards the need for EMS organisations in SA to improve their quality systems in general, if such measures are to be implemented [1.5, 1.6].

For the Rand method, the importance of having completed the practical data extraction using the case vignettes made a difference in the QI rating [2.1,2.2]. This expanded further into a general conversation about applying the QI framework, the quality system in which they'd be applied and documentation in

general [2.3 – 2.6].

Discussion

The simplicity and practicality of QIs as a system of quality measurement has led to their widespread adoption in healthcare[4,14,28–34]. Importantly, they align with Donabedian's conceptual framework for healthcare evaluation, predicated on the belief that an effective structure gives rise to effective processes of care, which in turn result in improved outcomes[8]. Within the PEC setting, patient exposure times are generally limited, and the delivery of care based largely around processes as opposed to outcomes. The utilisation of QIs as a measure of quality are therefore ideally suited to this environment.

Despite these advantages, the implementation of inappropriate or poorly tested QIs - even in well-established quality systems - has been reported to be both time-consuming and costly to correct[9,14]. Consequently, QI appraisal has been identified as an essential step toward understanding the appropriateness of these measures for a particular healthcare field or setting, prior to full implementation. The results of this study support these notions through the application of QI appraisal protocol against a series of QIs. Further to this, the results support the value in adopting a multi-method approach towards QI appraisal, compared to the single method approach.

From a series of 104 QIs identified for potential use in the SA PEC context, eight were identified as non-valid and three identified for which caution was recommended prior to full implementation, that were shared between the appraisal methods. A further 19 QIs were identified in the above categories and not shared by each method, highlighting the pragmatic advantages of a multi-method approach versus the single method approach. Our observations found the multi-method approach to be advantageous in that the methods complemented each other's strengths and compensated for each other's weaknesses. While the Qualify tool appraised the QIs from a greater number of viewpoints, the Rand approach offered insight into the practical application of the QIs not available with the Quality tool. This was additionally evident in the group discussion analysis, which in and of itself added further input towards understanding and appraising the appropriateness of the QIs that would not have otherwise been captured or understood by the categorical methods alone[18,35].

Despite these advantages, the application of the protocol required a significant investment in time and staff resources. The overall benefits of such an approach are therefore heavily dependant on the availability of these resources. This availability will likely vary significantly, depending on the quality system setting within which the protocol will be applied. These "system-focused" factors therefore have the potential to exert as much influence on the validity of the QIs as the setting in which the QIs will be implemented[36,37].

The outcomes of the appraisal have identified a significant number of QIs assessed to be valid and feasible for the SA PEC setting. The majority are centred around clinically focused processes of care, measures that are lacking in current performance assessment in EMS in SA. The importance and potential influence of the quality system in which the QIs will be implemented was further highlighted across all the methodologies. Quality system-focused assessment criteria, on average, scored lower than those criteria assessed to be characteristics inherent to the QIs themselves. This was reaffirmed during the qualitative discussion analysis, where system focused factors were a regular discussion point.

Conclusion

Measurement forms a central part of every healthcare quality system. Regardless of the measurement approach used, it is essential that the framework be comprehensively assessed for appropriateness for the setting in which it will be employed. Understanding and accounting for this as a factor is key towards ensuring both successful implementation and ongoing utilisation of such a system in this setting. The utilisation of a multi-method appraisal protocol offers significant benefit towards achieving this, when compared to the single appraisal approach, and can provide the confidence that the outcomes of the appraisal will ensure a strong foundation on which the measurement framework can be successfully implemented and employed.

Declarations

Ethics approval and consent to participate

Ethical approval for the study was granted by Stellenbosch University Health Research Ethics Committee (HREC) (Ref no. S15/09/193). Written consent for participation was provided by each of the participants prior to data collection. The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Consent for publication

Not applicable/required

Competing interests

The authors declare that they have no competing interests.

Funding

Nil

Authors' contributions

IH, PC, MC, LW and VL conceived the study. IH, conducted the data collection and analysis. IH drafted the manuscript, and all authors contributed to its revision. All authors have read and approve of the final manuscript, and consent to its publication. IH takes responsibility for the paper.

Acknowledgements

Nil

Abbreviations

PEC: Prehospital Emergency Care

QI: Quality Indicator

SA: South Africa

CPG: Clinical Practice Guideline

IHI: Institute for Healthcare Improvement

EMS: Emergency Medical Services

ECP: Emergency Care Practitioner

ACS: Acute Coronary Syndrome

CI: Confidence Interval

CINAHL: Cumulative Index to Nursing and Allied Health Literature

References

1. National Academies of Sciences; Engineering and Medicine. Crossing the global quality chasm: Improving health care worldwide. Washington, DC: The National Academies Press; 2018.
2. Mant J. Process versus outcome indicators in the assessment of quality of health care. Int J Qual Heal Care 2001;13(6):475–80.
3. Mainz J. Defining and classifying clinical indicators for quality improvement. Int J Qual Heal care 2003;15(6):523–30.
4. Mainz JAN. Developing evidence-based clinical indicators : a state of the art. Int Jounral Qual Healthc 2003;15(Suppliment 1):5–12.
5. Maio RF, Garrison HG, Spaire DW, et al. Emergency medical services outcomes project I (EMSOP I): Prioritizing conditions for outcomes research. Ann Emerg Med 1999;33(4):423–32.
6. Spaite DW, Maio RF, Garrison HG, et al. Emergency Medical Services Outcomes Project (EMSOP) II : Developing the Foundation and Conceptual Models for Out-of-Hospital Outcomes. Ann Emerg Med 2001;37(6):657–63.
7. Howard I, Cameron P, Wallis L, Castren M, Lindstrom V. Quality Indicators for Evaluating Prehospital Emergency Care: A Scoping Review. 2017;
8. Donabedian A. Evaluating the Quality of Medical Care. Milbank Q 1966;44(Suppl):166–206.
9. Campbell S, Kontopantelis E, Hannon K, Burke M, Barber A, Lester H. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. BMC Fam Pract. 2011;12:85.
10. BQS - INSTITUTE OF QUALITY AND PATIENT SAFETY. QUALIFY: Instrument for the Assessment of Quality Indicators. 2007;(August):60.
11. De Cruppé W, Kleudgen S, Diel F, Burgdorf F, Geraedts M. Feasibility of 48 quality indicators in ambulatory care in Germany: A cross-sectional observational study. Z Evid Fortbild Qual Gesundhwes. 2015;109(9–10):682–94.
12. Bieler D, Hörster A, Lefering R, et al. Evaluation of new quality indicators for the TraumaRegister DGU ® using the systematic QUALIFY methodology. Eur J Trauma Emerg Surg [Internet] 2018
13. Fitch K, Bernstein SJ, Aguilar MD, et al. The RAND/UCLA Appropriateness Method User's Manual. 2001;
14. Marshall Martin MN, Campbell Sue Kirk SM, Reeves D, Brook Associate Director RAND Health Elizabeth McGlynn Senior Scientist Paul G Shekelle RA. Measuring General Practice: A demonstration project to develop and test a set of primary care clinical quality indicators Professors of General Practice. 2003
15. Santori G, Fontana I, Valente R, Ghirelli R, Valente U. Application of the RAND/UCLA Appropriateness Method to Evaluate an Information System for Kidney/Pancreas Transplantation in Adult Recipients. Transplant Proc 2008;40(6):2021–3.
16. Bobrovitz N, Parrilla JS, Santana M, Straus SE, Stelfox HT. A qualitative analysis of a consensus process to develop quality indicators of injury care. Implement Sci. 2013;8:45.
17. Grudzen CR, Liddicoat R, Hoffman JR, Koenig W, Lorenz KA, Asch SM. Developing quality indicators for the appropriateness of resuscitation in prehospital atraumatic cardiac arrest. Prehospital Emerg care. 2007;11(4):434–42.

18. Curry LA, Nembhard IM, Bradley EH. Qualitative and mixed methods provide unique contributions to outcomes research. Circulation 2009;119(10):1442–52.
19. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. BMJ 2010;341(7783):1147–50.
20. Johnson RB, Onwuegbuzie AJ. Mixed Methods Research: A Research Paradigm Whose Time Has Come. Educ Res 2004;33(7):14–26.
21. Howick J, Phillips B, Ball C, Sackett D, Badenoch D. Oxford Centre for Evidence-based Medicine Levels of Evidence. 2009;Available from: https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/
22. Department of Health KZN. 2017 / 18 Annual Report. Annu. Rep. 2016/17. 2018;1–417.
23. North West Department of Health. Annual Report. 2018;
24. Western Cape Department of Health. Western Cape Government,Department of Health Annual Report. 2017;
25. Limpopo Department of Health. Annual Report. Annu. Rep. 2017/18 Financ. Year. 2018;
26. Howard I, Cameron P, Wallis L, Castrén M, Lindström V. Identifying quality indicators for prehospital emergency care services in the low to middle income setting: The South African perspective. African J Emerg Med. 2019;(April)
27. Hseih H-F, Shannon SE. Three Approaches to Qualitative Content Analysis. Qual Health Res 2005;15(9):1277–88.
28. Kollberg B, Elg M, Lindmark J. Design and implementation of a performance measurement system in Swedish health care services: a multiple case study of 6 development teams. Qual. 2005;14(2):95–111.
29. Coleman P, Nicholl J. Consensus methods to identify a set of potential performance indicators for systems of emergency and urgent care. J Health Serv Res Policy. 2010;15 Suppl 2:12–8.
30. Hoogervorst EM, van Beeck EF, Goslings JC, Bezemer PD, Bierens JJLM. Developing process guidelines for trauma care in the Netherlands for severely injured patients: results from a Delphi study. BMC Health Serv Res. 2013;13(1):79.
31. Abrams H, Belanger F, Carew D, et al. Prioritizing performance measurement for emergency department care : consensus on evidence- based quality of care indicators Steering committee. 2011;13(5):28–43.
32. Nakayama DK, Saitz EW, Gardner MJ, Kompare E, Guzik E, Rowe MI. Quality assessment in the pediatric trauma care system. J Pediatr Surg. 1989;24(2):159–62.
33. Mackenbach JP, McKee M. A comparative analysis of health policy performance in 43 European countries. Eur J Public Health. 2013;23(2):195–201.
34. Groene O, Klazinga N, Kazandjian V, Lombrail P, Bartels P. The World Health Organization Performance Assessment Tool for quality improvement in hospitals (PATH): An analysis of the pilot implementation in 37 hospitals. Int J Qual Heal Care 2008;20(3):155–61.
35. Van Harten WH, Casparie TF, Fisscher OAM. Methodological considerations on the assessment of the implementation of quality management systems. Health Policy (New York) 2000;54(3):187–200.
36. Kaplan HC, Provost LP, Froehle CM, Margolis PA. The model for understanding success in quality (MUSIQ): Building a theory of context in healthcare quality improvement. BMJ Qual Saf 2012;21(1):13–20.
37. Vaughn VM, Saint S, Krein SL, et al. Characteristics of healthcare organisations struggling to improve quality: Results from a systematic review of qualitative studies. BMJ Qual Saf 2019;28(1):74–84.

Quality Indicator Evidence Base Review

1. **Norris RM. A new performance indicator for acute myocardial infarction. Heart. 2001;85(4):395–401.**
2. **Sobo E, Andriese S, Stroup C, Morgan D, Kurtin P. Developing Indicators for EMS Evaluation and Quality Improvement. Jt Comm J Qual Improv. 1992;27(3):138–54.**
3. **Siriwardena a N, Shaw D, Donohoe R, Black S, Stephenson J. Development and pilot of clinical performance indicators for English ambulance services. Emerg Med J. 2010;27(4):327–31.**
4. **Siriwardena AN, Shaw D, Essam N, et al. The effect of a national quality improvement collaborative on prehospital care for acute myocardial infarction and stroke in England. Implement Sci. 2014;9:17.**
5. **Myers JB, Slovis CM, Eckstein M, et al. EVIDENCE-BASED PERFORMANCE MEASURES FOR EMERGENCY MEDICAL SERVICES SYSTEMS-AMODEL FOR EXPANDED EMS BENCHMARKING. Prehospital Emerg Care 2008;12:141–51.**
6. **Colwell C, Mehler P, Harper J, Cassell L, Vazquez J, Sabel A. Measuring quality in the prehospital care of chest pain patients. Prehospital Emerg Care. 2009;13:237–40.**
7. **Rosengart MR, Nathens AB, Schiff M a. The identification of criteria to evaluate prehospital trauma care using the Delphi technique. J Trauma 2007;62(March):708–13.**

8. Mohan D, Rosengart MR, Farris C, Fischhoff B, Angus DC, Barnato AE. Sources of non-compliance with clinical practice guidelines in trauma triage: a decision science study. Implement Sci. 2012;7:103.
9. HT S, Bobranska-Artiuch B, Nathens A, SE S. A systematic review of quality indicators for evaluating pediatric trauma care. Crit Care Med. 2010;38(4):1187-1196 10p.
10. Stelfox HT, Bobranska-Artiuch B, Nathens A, Straus SE. A systematic review of quality indicators for evaluating pediatric trauma care. Crit Care Med. 2010;38(4):1187–96.
11. Myers JB, Slovis CM, Eckstein M, et al. Evidence-Based Performance Measures for Emergency Medical Services Systems: A Model for Expanded EMS Benchmarking. Prehospital Emerg Care.2008;12(2):141–51.
12. Santana MJ, Stelfox HT. Development and evaluation of evidence-informed quality indicators for adult injury care. Ann Surg. 2014;259(1):186–92.
13. Dunford J, Domeier RM, Blackwell T, et al. Performance Measurements in Emergency Medical Services. Prehospital Emerg Care. 2002;6(1):107–13.
14. Nakayama DK, Saitz EW, Gardner MJ, Kompare E, Guzik E, Rowe MI. Quality assessment in the pediatric trauma care system. J Pediatr Surg. 1989;24(2):159–62.
15. Patterson PD, Lave JR, Martin-Gill C, et al. Measuring Adverse Events in Helicopter Emergency Medical Services: Establishing Content Validity. Prehospital Emerg Care. 2014;18(1):35–45. A
16. Hoogervorst EM, van Beeck EF, Goslings JC, Bezemer PD, Bierens JJLM. Developing process guidelines for trauma care in the Netherlands for severely injured patients: results from a Delphi study. BMC Health Serv Res. 2013;13(1):79.
17. Greenberg MD, Garrison HG, Delbridge TR, et al. INDICATORS FOR OUT-OF-HOSPITAL EMERGENCY MEDICAL SERVICES : PERSPECTIVE. Prehospital Emerg Care. 1997;1:23–7.
18. Meara PO. A generic performance framework for ambulance services : an Australian health services perspective. 2012;3(3).
19. Stelfox HT, Straus SE, Nathens A, Bobranska-Artiuch B. Evidence for quality indicators to evaluate adult trauma care: a systematic review. Crit Care Med. 2011;39(4):846–59.
20. Valenzuela TD, Spaite DW, Meislin HW, Clark LL, Wright AL, Ewy GA. Emergency vehicle intervals versus collapse-to-CPR and collapse-to-defibrillation intervals: Monitoring emergency medical services system performance in sudden cardiac arrest. Ann Emerg Med. 1993;22(11):1678–83.
21. Grudzen CR, Liddicoat R, Hoffman JR, Koenig W, Lorenz KA, Asch SM. Developing quality indicators for the appropriateness of resuscitation in prehospital atraumatic cardiac arrest. Prehospital Emerg care. 2007;11(4):434–42.
22. Patterson PD, Lave JR, Martin-Gill C, et al. Measuring Adverse Events in Helicopter Emergency Medical Services: Establishing Content Validity. Prehosp Emerg Care. 2013
23. Daudelin DH, Kulick ER, Amore KD, Lutz JS, Barrientos MT, Foell K. The Massachusetts Emergency Medical Service Stroke Quality Improvement Collaborative 2009 - 2012. Prev Chronic Dis. 2013;10.
24. Siriwardena AN, Shaw D, Essam N, et al. The effect of a national quality improvement collaborative on prehospital care for acute myocardial infarction and stroke in England. Implement Sci. 2014;9(17):1–9.
25. Oostema JA, Nasiri M, Chassee T, Reeves MJ. The quality of prehospital ischemic stroke care: Compliance with guidelines and impact on in-hospital stroke response. J Stroke Cerebrovasc Dis. 2014;23(10):2773–9. 0
26. Stelfox HT, Straus SE. Measuring quality of care: Considering measurement frameworks and needs assessment to guide quality indicator development. J Clin Epidemiol. 2013;66(12):1320–7.
27. Nakayama DK, Gardner MJ, Waggoner T. Audit Filters in Quality Assurance in Pediatric Trauma Care. J Pediatr Surg. 1993;28(1):19–25.
28. Patterson PD, Probst JC, Moore CG. Expected annual emergency miles per ambulance: an indicator for measuring availability of Emergency Medical Services resources. J Rural Health.

2006;22(2):102–11.

29. Greenberg MD, Garrison HG, Delbridge TR, et al. Quality indicators for out-of-hospital emergency medical services: the paramedics' perspective. Prehosp Emerg Care. 1997;1(1):23–7.
30. Gitelman V, Auerbach K, Doveh E. Development of road safety performance indicators for trauma management in Europe. Accid Anal Prev. 2013;60:412–23.
31. Willis CD, Gabbe BJ, Cameron PA. Measuring quality in trauma care. Injury 2007;38(5):527–37.

Tables

Table 1: Quality Indicator appraisal tool categories and criteria*

Category	No.	Subcategory Criterion
Relevance	R1	Significance: "The indicator covers aspects of quality of life, morbidity, or mortality."
	R2	Benefit: "Use of the indicator can have a positive effect on the quality of care."
	R3	Potential risks/side effects: "No risks are known/assumed which may result from the use of the indicator."
Scientific soundness	S1	Unambiguity of definitions: "The indicator is defined clearly and unambiguously."
	S2	Reliability: "It is a reliable measurement."
	S3	Risk adjustment: "The indicator is sufficiently adjusted to risk" (Are all factors that are not caused by the user taken into due account?)
	S4	Sensitivity: "The indicator provides sufficient sensitivity."
	S5	Specificity: "The indicator provides sufficient specificity."
	S6	Validity: "The indicator provides sufficient validity."
Feasibility	F1	Understandability and interpretability for patients and interested public
	F2	Understandability and interpretability for medical and nursing personnel
	F3	Possibility to influence the indicator manifestation: "The quality indicator refers to an aspect of care which can be influenced by the actors to be assessed."
	F4	Availability of data: "The data are documented by the service provider as a routine or can be collected with acceptable effort."
	F5	Data collection effort: "There is no data collection method available that provides at least equivalent results with less effort."
	F6	Implementation barriers: "Implementation barriers are unknown or covered by adequate measures."
	F7	Accuracy: "The correctness of the data can be verified."
	F8	Data integrity: "Is the individual data set intact?"
	F9	Completeness of the data: "Is it possible to verify that all occurring cases were recorded?"

*BQS – Institute of Quality and Patient Safety. QUALIFY: Instrument for the Assessment of Quality Indicators. 2007;(August)

Table 4: Literature review of evidence base

Indicator Category	Indicator subcategory	Total QIs	Indicator Type				Level of Evidence										Ref
			Structure	Process	Outcome	Sentinel Event	1a	1b	1c	2a	2b	2c	3a	3b	4	5	
Clinical	Acute Coronary Syndromes	25		23	2					4	5				2	14	1–6
	Airway management	8		8							2		1		1	2	2,7–11
	Acute Pulmonary Oedema	2		2						2							5
	Asthma	10		10						1						9	2,3,11
	General	18		15	3					2					4	12	2,6–9,12–19
	Hypoglycaemia	3		3												3	3
	Out of hospital cardiac arrest	44	4	38	2					2					3	39	2,3,5,7–9,13,18,20–22
	Pain management	1		1												1	12
	Seizures	2		2						2							11
	Stroke	11		11							3					8	3,23–25
	Trauma	16	3	11	2					4			5			6	2,5,9,12,19,26
Non-clinical	Adverse Event	25				25					9			11		5	7,8,10,14,15,19,27
	Deployable resources	15	13	2	2										5	13	18,28,29
	Dispatch/Call times	90	7	73	6					3	1		26	17	4	39	2,7–9,12,15,16,19,25,30
	Documentation	16	3	13								2		2	3	11	7–9,12,15,18,19,31
	Employee focused	16	16											2	2	12	7–9,13,18,29
	Service user rating/satisfaction	9		6	3										1	8	13,18,29
Total		311	46	218	20	25	0	0	0	20	20	2	32	32	25	182	
%			15%	70%	6%	8%	0%	0%	0%	6%	6%	1%	10%	10%	8%	59%	

2a. Systematic review of 2b and better studies

2b. Retrospective cohort study or prospective cohort with poor follow-up/low quality RCT

Table 2: Inter-rater reliability analysis of individual appraisal by the Quality Indicator Appraisal Working Group

	Methodology	% agreement [p value (95% Confidence interval)]			Kappa [p value (95% Confidence interval)]		
	Quality Indicator Appraisal Tool						
Relevance							
R1	Significance	90%	[<0.001	(0.8675 - 0.9350)]	0.90	[<0.001	(0.8587 - 0.9334)]
R2	Benefit	83%	[<0.001	(0.7934 - 0.8746)]	0.82	[<0.001	(0.7704 - 0.8669)]
R3	Potential risks/side effects	41%	[<0.001	(0.3887 - 0.4395)]	0.25	[<0.001	(0.2065 - 0.2840)]
Scientific Soundness							
S1	Unambiguity of definitions	81%	[<0.001	(0.7818 - 0.8465)]	0.80	[<0.001	(0.7664 - 0.8390)]
S2	Reliability	49%	[<0.001	(0.4614 - 0.5181)]	0.30	[<0.001	(0.2647 - 0.3434)]
S3	Risk adjustment	71%	[<0.001	(0.6789 - 0.7340)]	0.66	[<0.001	(0.6248 - 0.6975)]
S4	Sensitivity	80%	[<0.001	(0.7695 - 0.8395)]	0.78	[<0.001	(0.7426 - 0.8269)]
S5	Specificity	88%	[<0.001	(0.8502 - 0.9126)]	0.87	[<0.001	(0.8395 - 0.9093)]
S6	Validity	100%		(1)	1.00		(1)
Feasibility							
F1	Understandability and interpretability for patients and interested public	28%	[<0.001	(0.2670 - 0.2959)]	0.09	[<0.001	(0.0646 - 0.1076)]
F2	Understandability and interpretability for medical and nursing personnel	100%		(1)	1.00		(1)
F3	Possibility to influence the indicator manifestation	45%	[<0.001	(0.4286 - 0.4714)]	0.35	[<0.001	(0.3233 - 0.3835)]
F4	Availability of data	65%	[<0.001	(0.6434 - 0.6630)]	0.48	[<0.001	(0.4487 - 0.5134)]
F5	Data collection effort	22%	[<0.001	(0.2104 - 0.2345)]	0.01	[<0.001	(-0.0133 - 0.0235)]
F6	Implementation barriers	49%	[<0.001	(0.4803 - 0.5069)]	0.11	[<0.001	(0.0775 - 0.1503)]
F7	Accuracy	49%	[<0.001	(0.4803 - 0.5069)]	0.11	[<0.001	(0.0775 - 0.1503)]
F8	Data integrity	49%	[<0.001	(0.4765 - 0.5030)]	0.35	[<0.001	(0.3283 - 0.3695)]
F9	Completeness of the data	49%	[<0.001	(0.4765 - 0.5030)]	0.35	[<0.001	(0.3283 - 0.3695)]
	RAND method						
Clarity		85%	[<0.001	(0.8079 - 0.8854)]	0.83	[<0.001	(0.7865 - 0.8786)]
Necessity		48%	[<0.001	(0.4663 - 0.5033)]	0.39	[<0.001	(0.3663 - 0.4196)]
Acceptability		90%	[<0.001	(0.8682 - 0.9363)]	0.90	[<0.001	(0.8585 - 0.9347)]
Technical Feasibility		47%	[<0.001	(0.4401 - 0.4958)]	0.32	[<0.001	(0.2735 - 0.3568)]

2c. "Outcomes" Research; Ecological studies
3a. Systematic review of 3b and better studies
3b. Non-consecutive cohort study/Individual case control study
4. Case series
5. Expert opinion

Figures

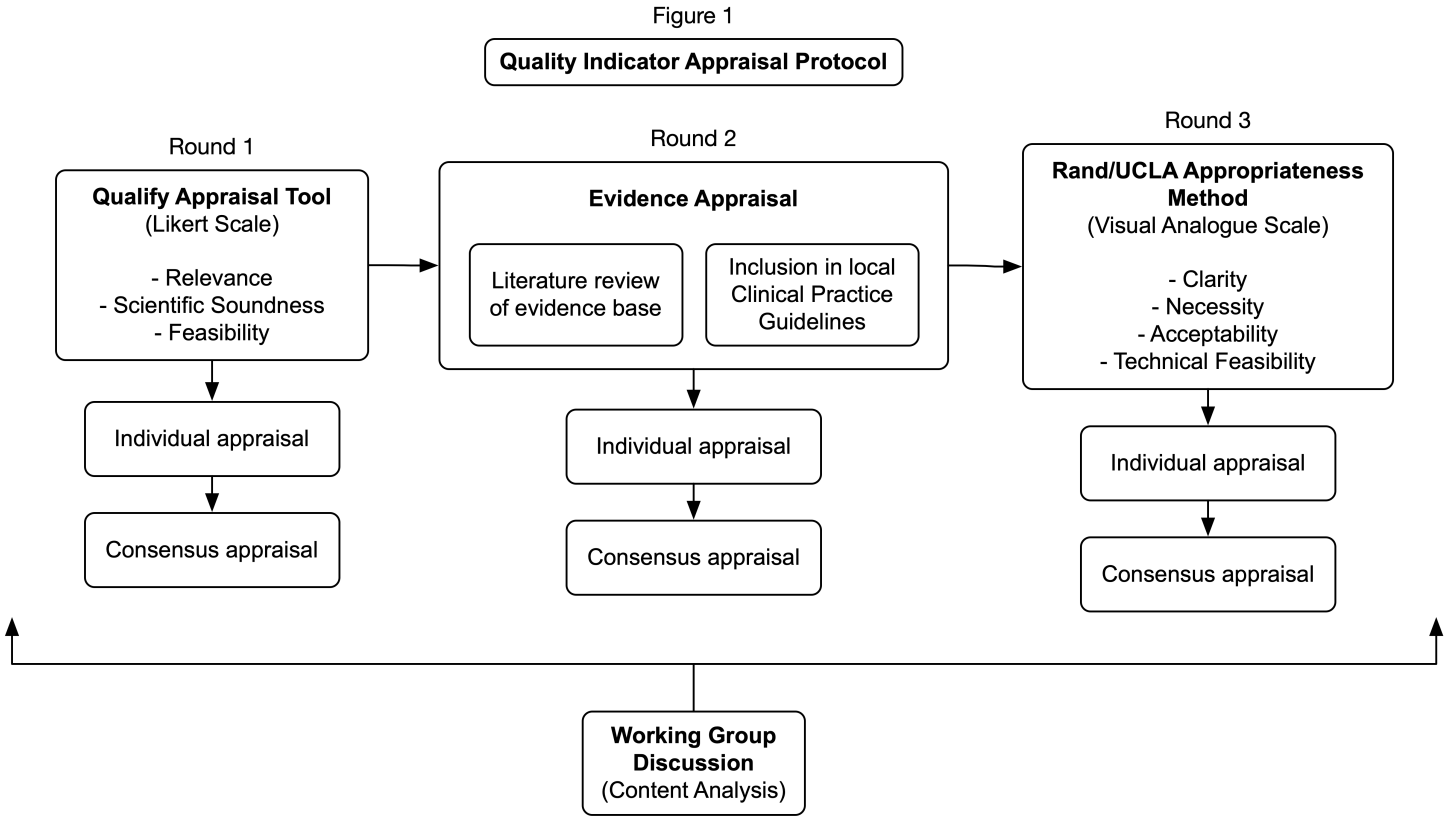

Figure 1

Quality Indicator Appraisal Protocol

Table 3: Quality Indicator appraisal results

Quality Indicator for Review	QI Class	Relevance	Scientific Soundness	Feasibility	Appraisal Tool Score	Total criteria Applies	Applicable CPG	Supported in CPG	Clarity	Necessity	Acceptability	Technical Feasibility	I
ACS/STEMI													
Patients with a provisional diagnosis of ACS/STEMI who had an ALS practitioner in attendance	Process	3.7	3.8	2.5	3.1	12	Yes	No	9.0	7.0	9.0	9.0	
Patients with a provisional diagnosis of ACS/STEMI who had a set of defined cardiac risk factors assessed and recorded	Process	3.3	3.8	3.1	3.4	16	Yes	Yes	5.0	5.0	7.0	4.0	
Patients with a provisional diagnosis of ACS/STEMI who had a 12 lead ECG obtained	Process	3.7	4.0	2.4	3.1	10	Yes	Yes	9.0	6.0	6.0	6.0	
Patients with a provisional diagnosis of ACS/STEMI who were administered Aspirin	Process	3.7	4.0	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients with a provisional diagnosis of ACS/STEMI who were administered GTN	Process	3.7	4.0	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients with a provisional diagnosis of ACS/STEMI who were assessed for suitability for thrombolysis by defined checklist	Process	3.7	3.8	1.8	2.8	10	Yes	Yes	5.0	5.0	7.0	4.0	
Patients with a provisional diagnosis of ACS/STEMI who were administered prehospital thrombolysis	Process	3.7	3.8	1.8	2.8	10	Yes	Yes	5.0	5.0	7.0	4.0	
Patients with a provisional diagnosis of ACS/STEMI who were transported directly to a Facility with PCI capabilities	Process	3.3	4.0	1.8	2.8	9	Yes	Yes	5.0	5.0	7.0	4.0	
Patients with a provisional diagnosis of ACS/STEMI who had EMS activation of the receiving Cath Lab	Process	3.7	3.8	1.8	2.8	10	Yes	Yes	5.0	1.0	1.0	4.0	
Patients who received/met all components of a defined ACS/STEMI composite bundle score	Process	3.7	3.8	3.1	3.5	15	No	No	7.0	7.0	8.0	6.0	
Acute Pulmonary Oedema													
Patients with a provisional diagnosis of APO who were administered GTN	Process	3.7	4.0	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients with a provisional diagnosis of APO who received CPAP	Process	3.7	4.0	2.6	3.3	11	Yes	Yes	9.0	9.0	9.0	2.0	
Patients with a provisional diagnosis of APO who had a 12 lead ECG obtained	Process	3.7	4.0	2.5	3.2	11	Yes	Yes	9.0	5.0	7.0	4.0	
Airway Management													
Patients who received a pre-ETI paralytic, following which there was a decrease in SpO2 > 10% from baseline/or decrease below 70% overall	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	7.0	9.0	9.0	9.0	
Patients successfully intubated by EMS personnel where EtCO2 monitoring was used post ETI	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients successfully intubated via RSI by EMS personnel where a paralytic agent was administered post-ETI	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients successfully intubated by EMS personnel where a sedative agent was administered post-ETI	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients successfully intubated by EMS personnel where a mechanical ventilator was used post-ETI for ventilation	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients in whom ETI was attempted by EMS personnel who had an alternative airway inserted as a final airway	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	8.0	9.0	9.0	9.0	
Patients in whom ETI was attempted by EMS personnel who had a surgical airway inserted	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients successfully intubated by EMS personnel with an EtCO2 < 30 mmHg or > 50 mmHg post-ETI > 10 mins during EMS care	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	7.0	9.0	9.0	9.0	
Patients in whom RSI with ETI was unsuccessful when attempted by EMS personnel	Process	3.7	3.8	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients in whom Non-RSI ETI was unsuccessful when attempted by EMS personnel	Process	3.7	3.8	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0	
Patients in whom RSI with ETI was	Process	3.7	3.8	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0	

successful when attempted by EMS personnel												
Total number of patients successfully intubated via RSI by EMS personnel	Process	3.7	3.8	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients who received/met all components of the defined Airway management composite Bundle score	Process	3.7	3.8	3.1	3.5	15	No	No	7.0	8.0	9.0	7.0
Anaphylaxis												
Patients with a provisional diagnosis of Anaphylaxis and evidence of bronchoconstriction documented who were administered a B2 agonist	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Anaphylaxis and evidence of bronchoconstriction documented who were administered an Anti-cholinergic bronchodilator	Process	3.7	3.7	3.0	3.3	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Anaphylaxis who were administered an antihistamine	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	7.0	9.0	9.0
Patients with a provisional diagnosis of Anaphylaxis who were administered a corticosteroid	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	6.0	9.0	9.0
Patients with a provisional diagnosis of Anaphylaxis and signs of a severe systemic response recorded who were administered IM Adrenaline	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	6.0	9.0	9.0	9.0
Asthma/Bronchoconstriction												
Patients with a provisional diagnosis of Asthma/Bronchoconstriction with lung sounds assessed and documented (pre and post treatment)	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Asthma/Bronchoconstriction with a SpO2 documented (pre and post treatment)	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Asthma/Bronchoconstriction who were administered a B2 agonist bronchodilator	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Asthma/Bronchoconstriction who were administered an anticholinergic bronchodilator	Process	3.7	3.7	3.0	3.3	15	Yes	Yes	8.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Asthma/Bronchoconstriction who were administered a corticosteroid	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Asthma/Bronchoconstriction recorded with documented severe wheezes/silent chest/BP < 90 mmHg systolic BP who were administered IM Adrenalin	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	7.0	9.0	9.0	9.0
Patients who received/met all components of the defined Asthma/Bronchoconstriction composite bundle score	Process	3.7	3.8	3.1	3.5	15	No	No	9.0	9.0	9.0	9.0
Burns												
Patients with a provisional diagnosis of Burns with burns dressings applied	Process	3.3	3.8	3.0	3.3	14	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Burns with body surface area and burns type assessed and recorded	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
General												
Serviceable suction unit devices available per defined area and/or time period	Structure	3.7	3.8	3.1	3.5	16	No	No	7.0	8.0	8.0	5.0
Serviceable 3 lead ECG monitoring devices available per defined area and/or time period	Structure	3.7	3.8	3.1	3.5	16	No	No	7.0	8.0	8.0	5.0
Serviceable 12 lead ECG monitoring devices available per defined area and/or time period	Structure	3.7	3.8	2.5	3.1	12	No	No	7.0	8.0	8.0	5.0

Serviceable portable oxygen cylinders available per defined area and/or time period	Structure	3.7	3.8	3.1	3.5	16	No	No	7.0	8.0	8.0	5.0
Serviceable Defibrillator/AED devices available per defined area and/or time period	Structure	3.7	3.8	3.1	3.5	16	No	No	7.0	8.0	8.0	5.0
Serviceable mechanical ventilators available per defined area and/or time period	Structure	3.7	3.8	2.5	3.1	12	No	No	7.0	8.0	8.0	5.0
Patients with reduced level of consciousness with a blood glucose measured	Process	3.7	3.8	3.1	3.5	16	Yes	Yes	9.0	9.0	8.0	9.0
Patients with a recorded SpO2 < 95% who were administered supplemental Oxygen	Process	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	9.0	8.0	9.0
Patients with a provisional diagnosis recorded	Process	3.7	3.8	3.0	3.4	15	No	No	9.0	9.0	9.0	9.0
Hypoglycaemia												
Patients with a blood glucose level < 5 mmol who were administered Glucose	Process	3.7	4.0	3.1	3.5	16	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a blood glucose level measured and recorded following Glucose administration	Process	3.7	4.0	3.1	3.5	16	Yes	Yes	9.0	9.0	9.0	9.0
Neonate/Paediatric												
One min APGAR score assessed and recorded for newborn patients	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Five min APGAR score assessed and recorded for newborn patients	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Paediatric patients with a provisional diagnosis of Croup who were administered oral/inhaled steroids	Process	3.7	4.0	3.0	3.4	15	Yes	No	9.0	9.0	9.0	9.0
Paediatric patients with a provisional diagnosis of Croup who were administered nebulised Adrenalin	Process	3.7	3.8	3.0	3.4	15	Yes	No	9.0	9.0	9.0	9.0
Patient transportation to a facility with specialist Paediatric capabilities/resources	Process	3.3	3.7	2.3	2.9	8	Yes	Yes	7.0	9.0	9.0	7.0
Obstetrics												
Obstetric patients who deliver prior to EMS arrival	Process	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	6.0	8.0	9.0
Obstetric patients with postpartum haemorrhage who were administered TXA	Process	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Obstetric patients with a provisional diagnosis of Eclampsia or Pre-eclampsia who were administered Magnesium sulphate	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Obstetric patients who deliver during EMS care	Outcome	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	8.0	9.0	9.0
OHCA												
Patients with a provisional diagnosis of OHCA with a witnessed collapse documented	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA who received documented bystander CPR	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA who received documented telephonic CPR advice	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	7.0	9.0	9.0	3.0
Patients with a provisional diagnosis of OHCA with VF/VT as first presenting rhythm on arrival of EMS	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA with Asystole/PEA as first presenting rhythm on arrival of EMS	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA intubated with alternative airway device	Process	3.7	3.8	2.4	3.1	10	Yes	Yes	8.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA for whom resuscitation was cancelled prior to arrival at hospital	Process	3.7	3.8	3.1	3.5	15	Yes	Yes	9.0	8.0	9.0	9.0
Patients with a provisional diagnosis of OHCA who were transported to hospital (incl. ROSC and Non-ROSC patients)	Process	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA with ROSC at hospital handover	Process	3.3	4.0	2.9	3.3	13	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional	Process	3.3	4.0	2.9	3.3	13	Yes	Yes	9.0	9.0	9.0	9.0

diagnosis of OHCA with VF/VT at hospital handover												
Patients with a provisional diagnosis of OHCA with Asystole/PEA at hospital handover	Process	3.3	4.0	3.0	3.4	14	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of OHCA with survival to Emergency Centre discharge	Process	3.3	3.8	1.6	2.6	8	Yes	Yes	7.0	9.0	9.0	2.0
Patients with a provisional diagnosis of OHCA with survival to hospital discharge	Outcome	3.3	3.8	1.6	2.6	8	Yes	Yes	7.0	9.0	9.0	2.0
Pain Management												
Patients with level of Pain measured via defined pain score	Process	3.7	3.8	3.1	3.5	16	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a defined pain score threshold who were administered analgesia	Process	4.0	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with level of pain measured via defined pain score following analgesia administration	Process	4.0	3.8	3.1	3.5	16	Yes	Yes	9.0	9.0	9.0	9.0
Seizures												
Patients with a provisional diagnosis of Seizures with a blood glucose measured and recorded	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Seizures who were administered an antiepileptic for ongoing Seizures	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Stroke/TIA												
Patients with a provisional diagnosis of Stroke/CVA/TIA with a blood glucose measured and recorded	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Stroke/CVA/TIA with a Stroke screening assessment performed (e.g.: FAST)	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	7.0
Patients with a provisional diagnosis of Stroke/CVA/TIA with serial blood pressure measurements recorded (X3)	Process	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients with a provisional diagnosis of Stroke/CVA/TIA delivered to a specialist Stroke Centre	Process	4.0	3.8	1.6	2.8	9	Yes	Yes	7.0	9.0	9.0	1.0
Patients with a provisional diagnosis of Stroke/CVA/TIA with direct delivery to CT scan	Process	3.3	3.8	2.3	3.0	8	Yes	Yes	7.0	9.0	9.0	2.0
Patients who received/met all components of the defined Stroke/CVA/TIA composite bundle score	Process	3.7	3.8	3.0	3.4	15	No	No	7.0	7.0	9.0	7.0
Trauma												
Patients designated as a trauma case with entrapment on scene documented	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients designated as a trauma case with a BP < 90 mmHg	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients designated as a trauma case with partial/full amputation who had a tourniquet applied	Process	3.7	4.0	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients designated as a trauma case with a femur fracture and traction splint use	Process	3.7	3.8	3.0	3.4	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients designated as a trauma case with a BP < 90 mmHg who were administered TXA	Process	3.7	3.8	3.1	3.5	15	Yes	Yes	9.0	9.0	9.0	9.0
Patients designated as a trauma case with direct transportation to a specialist Trauma Centre	Process	3.7	4.0	2.3	3.1	9	Yes	Yes	7.0	9.0	9.0	9.0
Adverse Events												
Number of patient deaths while in EMS care	Sentinel Event	3.7	4.0	2.4	3.1	10	No	No	7.0	9.0	9.0	8.0
Number of defined Adverse Events reported during EMS care	Sentinel Event	3.7	3.8	2.5	3.1	12	No	No	7.0	9.0	9.0	7.0
Number of defined equipment/technical failures reported during EMS care	Sentinel Event	3.7	3.8	2.5	3.1	12	No	No	7.0	9.0	9.0	6.0
Number of accidental or unexpected extubations reported during EMS care	Sentinel Event	3.7	3.8	2.5	3.1	12	No	No	8.0	9.0	9.0	4.0
Number of patients with a decrease in GCS of 3 or more points during EMS care	Sentinel Event	3.7	3.8	2.5	3.1	12	No	No	9.0	9.0	9.0	8.0
Number of defined failed intubation attempts	Sentinel Event	3.7	4.0	2.5	3.2	11	No	No	9.0	9.0	9.0	6.0
Total number of patient injury	Sentinel	3.7	3.8	2.5	3.1	12	No	No	7.0	9.0	9.0	7.0

reports during EMS care	Event										
Number of EMS staff on-duty injury reports	Sentinel Event	3.7	3.8	2.5	3.1	12	No	No	7.0	9.0	9.0	7.0
Number of defined medication errors during EMS care	Sentinel Event	3.7	3.8	2.5	3.1	12	No	No	7.0	9.0	9.0	6.0
					Communications/Dispatch							
Number of cases compliant with defined ALS Dispatch criteria	Structure	3.7	4.0	3.1	3.5	15	No	No	8.0	8.0	9.0	6.0
Number of cases with call processing time within defined limits	Structure	3.7	3.8	3.1	3.5	15	No	No	9.0	7.0	9.0	7.0
Number of Service Call Centre calls received per 10000 population	Structure	3.7	4.0	3.0	3.4	15	No	No	7.0	9.0	9.0	5.0
Number of unanswered/missed calls to the Service Call Centre	Structure	3.7	4.0	3.1	3.5	15	No	No	7.0	9.0	9.0	6.0
Number of cases with a delay in dispatch and/or response time waiting for a police/security escort	Process	3.7	4.0	2.6	3.3	11	No	No	8.0	9.0	9.0	7.0

Appraisal Tool Score < or on threshold of validity score
RAND Method Score < or on threshold of validity score

Table 5: Qualitative analysis of the Working Group discussion

Methodology	Text Reference	Sub-category	Supporting Quote
Quality Indicator appraisal tool	1.1	Relevance	"For me, because practically zero clinical indicators are used or reported publicly by EMS [Emergency Medical Services] in South Africa, their relevance wand significance and benefit was naturally going to be scored high"
	1.2	Usability	"Whenever I was rating a category that I used or drew information from the data dictionary, there was always sufficient information that left no doubt that it was well planned for or accounted for. The difficult part was knowing how much variation there would be in different EMS organizations in South Africa in how they would be able to extract this information and put it to use"
	1.3	Context	"Whatever indicators are used by a service, it's important that they do a feasibility assessment of what's possible for them to achieve. We may be able to say overall, like these will work for South Africa in general, but when it comes to actual implementation, a service is going to have to understand its surroundings and the types of patients it sees"
	1.4		"Like, the indicators involving direct transport to a CT [Computed Tomography] scanner for Stroke patients, or to PCI [Percutaneous Coronary Intervention] facilities for STEMI [ST Elevation Myocardial Infarction], those will only be applicable to certain metropolitan areas, and probably only for certain private services as well. It won't be a general indicator for everyone to use"
	1.5	Quality system	"This is a complete mind shift from what we currently know and how we measure quality in South Africa. If a service is serious about implementing these, even it's just a few, they're going to have to admit that it's going to take an overhaul in their quality system, and that it's likely going to need more resources than what they dedicate to measuring response times at the moment"
	1.6		"Outside of a few of the large private services, the provincial services are going to have to ramp up the effort around measuring quality. As simple and as easy a system that these indicators are, there's probably not many of the provincial services that are ready to implement them"
RAND method	2.1	Methodology	"You really get to see how these will be used from a practical point of view. I can see the benefit of how a simple system that's objective can make the world of difference. It's not like how I used to remember it when we checked the case sheets, and it depended on how you felt at the time"
	2.2		"Doing the data extraction made a big difference, because I remember, especially for the sentinel event indicators, I scored them quite low with the appraisal tool, but when we went through them and applied them to actual cases, it was much simpler than I thought it would be and so I scored them higher after being able to actual do the extraction"
	2.3	Technology	"I think applying these indicators would be way easier with an electronic patient report form. It's going to take way more effort in doing it manually, but I can still see the benefits even if it's done this way"
	2.4	Quality system	"I think when you're sitting down and applying the indicators to case sheets, the system does seem simple and straightforward enough to use. But what do you do from there? It's going to be a logistical challenge to get the paperwork together to do the assessment, but I feel like the bigger challenge is using the information we learn, it's just as important as getting the information"
	2.5	Transparency	"It seems like it's going to be easy to game the system. Like how I know the guys have done the things that they've written down. What sort of mechanism is there for to check that they've been truthful in their notes, especially if they now know they're being watched"
	2.6	Technology	"I think [participant] was right about the electronic record, because we can build checks and balances into that sort of thing to monitor truthfulness I suppose, also like [respondent] mentioned. That also solves the legibility issue and whether or not enough information has been written. Look at when we used the poor documentation examples, it was difficult to apply the indicators to those just because you didn't always have the right information to go on"

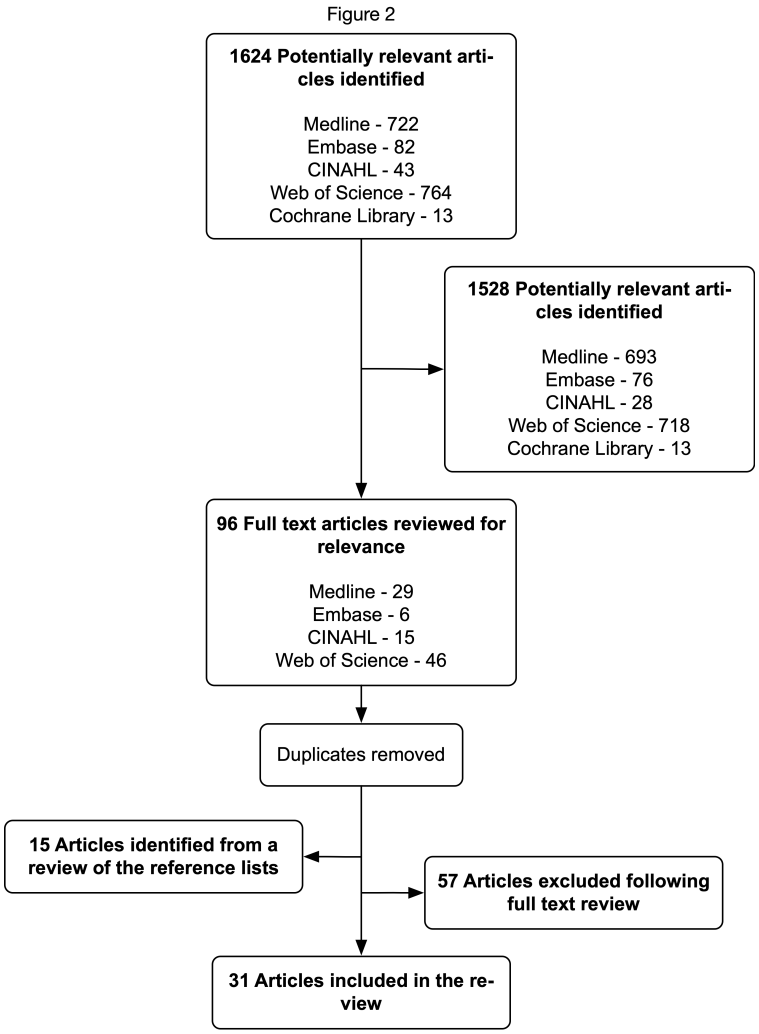

Figure 2

Selection of Articles for Review