Check for updates

RESEARCH ARTICLE

REVISED Development of performance indicators for systems of urgent and emergency care in the Republic of Ireland. Update of a systematic review and consensus development exercise [version 2; peer review: 2 approved, 2 approved with reservations]

Previously titled: Development of performance indicators for systems of urgent and emergency care in the Republic of Ireland: Systematic review and consensus development exercise

Siobhan Boyle[1], Rebecca Dennehy[1], Orla Healy[2], John Browne[1]

[1]School of Public Health, University College Cork, Cork, Ireland
[2]Department of Public Health, Health Service Executive, Dublin, Ireland

Abstract

Objectives: To develop a set of performance indicators to monitor the performance of emergency and urgent care systems in the Republic of Ireland.

Design: This study comprised of an update of a previously performed systematic review and a formal consensus development exercise. The literature search was conducted in PubMed and covered the period 2008 to 2014. The results of the review were used to inform a consensus group of 17 national experts on urgent and emergency care in Ireland. The consensus development exercise comprised an online survey followed by a face-to-face nominal group technique meeting. During this meeting participants had the opportunity to revise their preferences for different indicators after listening to the views of other group members. A final online survey was then used to confirm the preferences of participants.

Results: Initial literature searches yielded 2339 article titles. After further searches, sixty items were identified for full-text review. Following this review, fifty-seven were excluded. Three articles were identified for inclusion in the systematic review. These papers produced 42 unique indicators for consideration during the consensus development exercise. In total, 17 indicators had a median of greater than 7 following the meeting and met our pre-specified criterion for acceptable consensus.

Discussion: Using this systematic review and nominal group consensus development exercise, we have identified a set of 17 indicators, which a consensus of different experts regard as potentially good measures of the performance of urgent and emergency care systems in Ireland. Pragmatic implications are discussed with reference to three subsequently performed original studies which used some of the indicators

Open Peer Review

Reviewer Status ✓ ? ? ✓

	Invited Reviewers			
	1	**2**	**3**	**4**
REVISED **version 2** published 12 Feb 2019	✓ report		? report	✓ report
version 1 published 28 Feb 2018	? report	? report		

1 **Robbert Huijsman** (iD), Erasmus University Rotterdam, Rotterdam, The Netherlands

2 **Ellen Weber**, University of California, San Francisco, San Francisco, USA

3 **Maaret Castrèn**, Helsinki University Hospital and Helsinki University, Helsinki, Finland

4 **Anna Cantrell** (iD), University of Sheffield, Sheffield, UK

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: John Browne (j.browne@ucc.ie)

Introduction

Emergency and urgent care consists of all the services which contribute to the management of people when immediate care is sought for a health condition. When patients need immediate care they can enter the health system through a range of services and will often use more than one. This can lead to a duplication of services, confusion about the most appropriate access point for individual patients and the danger of poorly co-ordinated care, especially at the point where patients transfer from one service to another. Emergency and urgent care services include pharmacy, primary care, minor injury units, acute medical assessment units, emergency departments, mental health services and all the services required to refer and transport patients to an appropriate treatment facility.

There is an increasing awareness that urgent and emergency care services should operate as whole systems of care for the populations they serve[1]. Adopting such an approach requires individual service providers to be integrated into larger systems and to co-ordinate their activities accordingly. It is hoped that a systems approach will deliver a higher standard of quality, safety, efficiency, timeliness and overall patient experience without introducing inequity of access. Policy makers have a variety of tools at their disposal when attempting to engineer a systems approach to urgent and emergency care. These include the centralisation of care for high risk cases at high volume hospital units and the use of referral pathways and new facilities such as minor injury units to direct low-risk cases to settings that are appropriate for their condition[2]. Other elements include the use of telemedicine to provide support to smaller facilities and the development of community services for patients with conditions that are sensitive to the quality of ambulatory care[3].

The Health Service Executive (HSE) is responsible for the provision of publically funded health services in the Republic of Ireland. The HSE has attempted to foster a systems approach to urgent and emergency care services across the whole country, but the pace and nature of change is highly variable. In four peripheral regions (South, West, Mid-West, North-East) the reconfiguration process is at an advanced stage, but progress has been much slower in Dublin, the Midlands and the South-Eastern part of the country. This variation represents a natural experiment in policy making and is an opportunity to observe the impact of the changes that have been introduced before they are implemented across the whole country.

Existing indicators of urgent and emergency care performance focus on individual services and do not capture the performance of systems[4]. The development of such indicators would allow policy makers to compare different models of care and evaluate the longitudinal impact of changes to service configuration. In light of this and considering the introduction of a system-based approach to urgent and emergency care by the HSE, the aim of this study was to develop a set of performance indicators to monitor the performance of emergency and urgent care systems in the Republic of Ireland.

Methods
Systematic review
This study comprised of an update of a previously performed systematic review which had covered the period up to 2007[5] and a formal consensus development exercise. The systematic review update was conducted in August 2014 by one person (RD). Articles cited in PubMed over the period 2008 to 2014 were systematically searched by combining variations of the text terms 'emergency' and 'indicator' using the AND operator. Our search for novel indicators was supplemented by a review of the reference lists of articles selected for review. We also contacted experts and organisations working on the assessment of urgent and emergency care performance to identify relevant grey literature. These included the Society for Academic Emergency Medicine (USA), the Centre for Medicare and Medical Services (USA), the Emergency Department Benchmarking Alliance (USA), the Canadian Association of Emergency Physicians, the European Society for Emergency Medicine, the Royal College of Surgeons England, the Pre-Hospital Emergency Care Council (Ireland) and relevant HSE Clinical Programme Directorates.

Articles were selected for review by two persons (RD and JB) on the basis that they might contain definitions of system-level indicators of emergency and urgent care performance. Articles were excluded after review of the full text version if the indicators that they contained were already listed by the previous systematic review or if they focused on individual components of the urgent and emergency care system such as emergency

department waiting times or ambulance response times. Non-English articles were also excluded. Included articles were read and discussed by RD and JB who came to a consensus about whether they contained at least one indicator that measured system-level performance in urgent and emergency care. A conservative approach was taken and indicators were excluded only if they clearly related to a specific service and did not capture elements of system function. This approach was warranted on the basis that the subsequent consensus development exercise was to be the definitive judgement on the extent to which indicators were useful signals of system performance.

The systematic review has been reported according to PRISMA (Supplementary File 1).

Consensus development exercise

The consensus development exercise comprised an online survey and a face-to-face nominal group meeting. A broad range of Irish-based experts were recruited to the consensus development group. Experts were recruited by contacting professional representative bodies, policy making organisations, regulatory bodies and patient advocacy groups. The following clinical disciplines were recruited to the group: emergency nursing, acute medicine, minor injuries/urgent care nursing, anaesthesia/intensive care, emergency medicine, psychiatry, public health, paediatrics, pre-hospital care, general practice, pharmacy and geriatric medicine. The HSE quality improvement directorate, the Irish Department of Health, the Irish healthcare regulator (Health Information and Quality Authority) and two patient advocacy groups were also represented. Once individuals were highlighted as potential members, they were approached through email and phone calls to join the group. Interested parties were then sent a formal invitation letter to join the group. In total the group was composed of 17 national experts on urgent and emergency care in Ireland.

All novel indicators identified in the updated systematic review were combined with those identified in the original systematic review and grouped under the following headings in an online survey: outcome based indicators, process based indicators and structural indicators (see Supplementary File 2). The definitions of urgent and emergency conditions were adopted from those used in previous consensus development work[4] performed by the University of Sheffield for the English NHS (see Supplementary File 3). The survey was designed and distributed to the consensus development group using the online tool, Survey Monkey. All members of the group were sent a link to the online survey and asked to complete it. Each member was asked to rate their agreement with the statement 'this measure is likely to be a good indicator of the performance of the emergency and urgent care system', on a Likert scale anchored by 1 ('disagree strongly') and 9 ('agree strongly'). There was also space for members to add any comments. Participants were asked not to limit their views about the potential usefulness of an indicator by perceived difficulties in collecting or processing the data required to calculate them. 14 consensus group members completed the survey.

A face to face meeting was held in October 2014 and all members of the consensus group who had completed the online survey were asked to attend. Thirteen of the 14 invited members attended the meeting. One member had planned to attend but was not able to do so because of a separate work commitment which arose close to the meeting. Each participant was provided with the original questionnaire which now included a record of their individual responses to the online survey and the group's median score and interquartile ranges. The meeting was conducted using a nominal group technique format. Once each participant had been given the opportunity to provide their opinion about an indicator, that indicator was ranked again by the members of the group. This procedure was followed for each individual indicator until all indicators had been discussed. Following the meeting, the performance indicators were ranked by their median agreement score. Those with a median greater than 7 were classified as potentially good performance indicators. A second online survey was then created using the online tool, Survey Monkey. Those indicators which had scored a median greater than 7 were included and all participants were asked to rank these indicators in order of preference. Higher ranks indicate greater preference and are represented by lower numbers. This exercise was sent to the 13 members of the group who had attended the consensus development meeting and there was a 100% completion rate.

Results

The literature search strategy identified 2339 article titles. A title search reduced this to 150 articles and a review of the abstracts of these papers led to retrieval of 47 articles for a full-text review. A further seven articles were identified from the reference lists of the 47 full articles that were reviewed and six other documents from grey literature sources were selected for review. Two researchers reviewed the sixty items selected for full-text review (RD and JB). Following this review, fifty-seven were excluded for the following reasons: forty-four of the articles excluded at the full text stage were focused on service based indicators, seven reported on indicators that had been described by the previous systematic review and six were of a descriptive nature and not focused on specific indicators (Figure 1). This process led to three articles being identified for inclusion in the review[4,6,7].

The three articles included in the final review yielded four novel indicators that had not been presented in the previous systematic review. These were: patient reported experience of whole episodes of emergency and urgent care[4]; mortality rates among inter-hospital transfer patients[6]; inter-hospital transfer times[6]; and time from decision to admit to transfer to an appropriate inpatient bed[7]. The combination of these new indicators with those that had been identified in the previous systematic review produced 42 unique indicators for review by the consensus development group. In total, 17 indicators had a median score on the Likert scale of agreement of greater than 7 following the consensus meeting (see Supplementary File 4). Table 1 presents the median, mean and range of rankings for
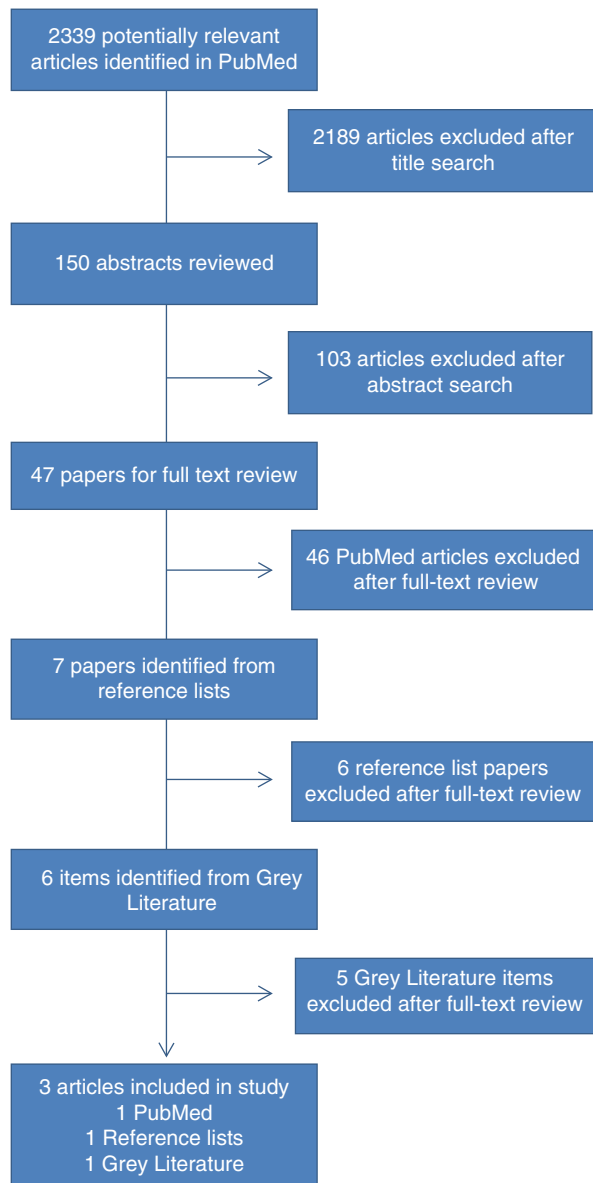
Figure 1. Flow chart of literature search.

these 17 indicators that were produced by the second online survey.

Discussion

Using a systematic review and nominal group consensus development exercise, we have identified a set of 17 indicators which a consensus group of different experts regard as potentially good measures of the performance of urgent and emergency care systems in Ireland. As far as we are aware this is only the second attempt to produce a set of indicators that are purely focused on system-level performance in urgent and emergency care and the first for Ireland. The indicator set was designed by Irish experts to be applicable to the Irish health system. Thus, it may not necessarily be applicable to health systems in other countries because of differences in terminology and how services are organised.

The list is made up of twelve process and five outcome indicators. Four of the seventeen indicators were included in the top sixteen indicators produced by a previous consensus development exercise carried out in the UK[4] and a further three were novel indicators which were identified through our systematic review. The three novel indicators were patient reported experience of whole episodes of emergency and urgent care[4]; mortality rates among inter-hospital transfer patients[6]; and time from decision to admit to transfer to an appropriate inpatient bed[7].

This study was undertaken using standard systematic review and consensus development methods. The members of the consensus group were purposively chosen as they were identified as having a wide range of expertise and knowledge in relation to various aspects of emergency and urgent care. The online survey allowed the opinions of those members to be collected and aggregated, while the face to face meeting offered the opportunity for the members to consider the indicators in light of hearing the opinion of their colleagues, as well as enabling discussion among panellists on the wording and clarity of the performance indicators.

Study limitations

Our study has some limitations. The systematic review only covers the period up to 2014 and may, therefore, miss some more recent literature. Since 2014 at least two further consensus development studies have been published in this field. These studies were not focused on the whole system of urgent and emergency care: one covered only prehospital care[8] and one focused only on care in the ED[9]. Nevertheless, it is notable that the study of prehospital performance found that "direct transport of ST-elevation myocardial infarction patients to a primary percutaneous intervention (PCI)-capable facility for ECG to PCI time <90 min" was the highest ranked indicator. This is consistent with our finding that 'time from call to care' for indicator conditions is rated as the best performance indicator for a whole system of urgent and emergency care. This provides some support to the validity of our consensus development exercise.

Data extraction in the systematic review was conducted independently by two reviewers who then came to consensus. However, no formal assessment of inter-rater reliability was conducted.

No attempt was made to achieve unanimity so it is possible that some of the indicators may be controversial to certain stakeholder groups. We also requested that panel members did not consider the feasibility of collecting data required to calculate an indicator. This may mean that the chosen performance indicators are not immediately measurable; however, we are hopeful that progress in data collection may allow these performance indicators to be measured in the future.

Table 1. Final ranking of 17 indicators likely to be useful indicators of the performance of Irish urgent and emergency care systems.

Median rank (range)	Mean rank	Performance indicator	Explanation
2 (1–10)*	3.9	Time from call to care for indicator conditions. E.g. for patients having thrombolysis, call to needle times. For patients having percutaneous coronary intervention (PCI), call to cath lab, for patients undergoing hip fracture repair, call to theatre	The aim of this indicator is to reduce times on patient journey through the EUCS to definitive care. It has been suggested that data for this indicator could be sourced from national and local recommendations for 'definitive care' for SEC and urgent conditions; service level linked data; ED data; AS patient report forms; and theatre books
3 (1–11)	4.4	Case fatality rates for serious, emergency conditions for which a well-performing EUCS could improve chances of survival	This indicator is based on health outcomes and it aims to reduce the proportion of patients with specified serious emergency conditions who die. This indicator could be calculated through HIPE and CSO Mortality Statistics.
7 (1–15)	7	Adherence to any evidence-based good practice guidelines for serious emergency, and urgent conditions	The aim of this indicator is to encourage services within EUCS to adopt good practice in managing patient care in accordance with the best available research evidence in published guidelines. This indicator would be measured through the auditing of practice and procedures that are implemented in EUCS.
7 (1–17)	7.5	Mortality rates among inter-hospital transfer patients for this group of conditions	This indicator aims to examine the best practice process of transfer of patients from one hospital to another and the mortality rates associated with this process. Data for this indicator may be collected by accessing both ambulance and HIPE data.
7 (3–16)	9.1	For EUCS users with the following group of serious, emergency conditions, who are admitted, the time from call to ambulance service to admission	This indicator aims to ensure that patients, who are admitted with serious emergency conditions, do so in an appropriate and timely manner. This data could be collected through HIPE, and ambulance data.
8 (2–16)	8.3	Call to ambulance service to time on scene.	This indicator aims to examine the variations in outcomes or processes due to differences in access and availability of care. A well performing EUCS will deliver or be working to deliver the same processes of care at all times and in all places. In order to measure this indicator, data sources required would include National Ambulance Service, GP records and Patient Surveys.
8 (1–17)	8.9	Patient reported experience of whole episodes of emergency and urgent care.	This can be measured through surveys similar to that which will be carried out as part of SIREN Work Package 4. The questionnaire addresses three domains of patient experience: entry into the system; progress through the system; and convenience of the system.
9 (2–15)	8.3	Time from decision to admit to transfer of patient to appropriate in-patient bed.	This indicator aims to ensure that patients who are admitted are appropriately placed in an in-patient bed in a timely manner. The data collected should include times of first contact, assessment and critical points in the patient's journey. It could be measured through patient surveys and hospital audits.
9 (1–15)	8.9	Emergency re-admissions within 28 days as a proportion of all live discharges for the following group of urgent conditions	This indicator focuses on the processes within the emergency and urgent care system and its aim is to encourage services to work collaboratively in order to manage care both in hospital and in the post-discharge period. Data for this indicator could be sources from HIPE
9 (3–16)	10.4	For all of the serious emergency conditions combined, the proportion of deaths that occur before admission (i.e. in pre-hospital or in the Emergency Department)	This indicator aims to examine those patients with serious emergency conditions who die before admission to either pre-hospital or an ED. Data could be obtained from GPs
10 (3–16)	9.6	Hospital emergency admission rates for the following group of urgent conditions whose exacerbations could be managed out of hospital or in ED's without admission to an inpatient bed	This focuses on avoidable admissions for acute exacerbations of urgent conditions. This indicator aims to reduce hospital admission rates for episodes that could be managed out of hospital or in settings without admission to a hospital bed.
11 (5–17)	10.9	Time from patient arrival at referring hospital to making the decision to transfer	This focuses on the processes associated with patient transfer to a hospital setting. It could be calculated using ambulance data

Median rank (range)	Mean rank	Performance indicator	Explanation
12 (1–17)	10.2	Case fatality rates for serious, emergency conditions for which a well-performing EUCS could improve chances of survival but for out of hospital deaths	This indicator is based on health outcomes and it aims to reduce the proportion of patients with specified serious emergency conditions who die outside of the hospital setting. Data could be obtained from GP's and the coroner's office.
12 (1–17)	10.6	Time from onset of serious emergency condition to arrival at the receiving hospital	This focuses on the processes associated with patient transfer to a hospital setting. It could be calculated through patient survey and ambulance data
12 (5–17)	12	Time from decision to transfer to availability of transport	This focuses on the processes associated with patient transfer to a hospital setting. It could be calculated using ambulance data
13 (4–17)	10.8	For EUCS users with the following group of serious, emergency conditions, who are admitted, the time from first contact with mental health team to admission	This indicator aims to ensure that patients, who are admitted with serious emergency conditions, do so in an appropriate and timely manner. This data could be collected through HIPE.
13 (4–17)	12.2	Arrivals at ED referred by any EUCS service and discharged without treatment or investigation(s) that needed hospital facilities	This indicator aims to enable EUCs to monitor the quality of the initial assessment and appropriate triage at the first point of contact. ED records could be used as the data source

*Higher ranks indicate greater preference and are represented by lower numbers.

Implications

The results of this paper have been used to conduct three studies on urgent and emergency care in Ireland, each using one of the highly ranked indicators described above. The first study measured "case fatality rates for serious, emergency conditions for which a well-performing EUCS could improve chances of survival" to assess whether national and regional population outcomes in Ireland over the period 2002–2014 were associated with the reconfiguration or regional urgent and emergency care systems. No distinct pattern of change was found among regions which underwent substantial reconfiguration compared to those that did not[10]. The second study measured the impact of population and health system factors on county-level variation in "conditions whose exacerbations could be managed out of hospital or in ED's without admission to an inpatient bed" in Ireland over the period 2014–2016[11]. This found that potentially avoidable emergency admissions are primarily driven by socio-economic conditions, hospital admission policies and private health insurance coverage and were not associated with primary care resources. This is an important finding for policy in this area as it suggests that reductions in emergency admissions for ambulatory conditions will not necessarily be achieved by simply increasing the quantity of relevant primary care resources. The study also found that the distinction between 'potentially avoidable' and all other emergency admissions may not be as useful as previously believed when attempting to identify the causes of regional variation in emergency admission rates. The third study used "patient reported experience of whole episodes of emergency and urgent care" to compare user experiences of eight regional urgent and emergency care systems in Ireland[12]. It found no consistent relationship between patient experience and the type of urgent and emergency care model in different regions, and concluded that composite questionnaire data may not offer a useful metric for exploring the impact of system-level service change.

Our research programme, including the present study, have produced some consistent learning points which may be useful for researchers in this field. First, it is not easy to distinguish between indicators of the performance of individual services and the system as a whole because of the inter-dependency of different services. For example, the indicator "call to ambulance time on scene" at first glance seems to cover only one service, but during the discussions of the consensus group it was felt to reflect a number of system-level issues such as the geographical configuration of ambulance stations and Emergency Department locations, and capacity pressures on the ambulance service that are caused by system-level decisions (e.g. need to spend more ambulance resources on inter-hospital transfers).

Second, it is difficult to operationalise research where populations, interventions and outcomes are difficult to define. For example, in all three of the studies described above the populations studied were compared at geographical levels (e.g. county, region) but these are inevitably an imperfect level of analysis because geographical units do not map perfectly to health system interventions, especially at boundaries where populations are exposed to both old and new models of care. This problem did not fatally undermine the internal validity of our case-fatality study[10] because all geographical units were subject to some 'boundary leakage' of populations, and the proportion of the overall study populations affected was low. In general, however, researchers should be sensitive to instances where this leakage is likely to pose a substantial challenge, for example when boundaries cut across urban areas and a high proportion of the population has easy access to more than one model of care.

A third, related issue, is the overall complexity of system evaluation. It is debatable whether traditional study designs are a useful way to compare whole systems of care because of the

difficulty in making strong inferences about causal relationships within complex systems. Although there are growing calls for evaluations of complex organisational interventions at the level of whole health systems, progress has been limited because of the difficulty in dealing with complexity[13]. In two of our studies[10,12] we found no evidence of difference in outcomes between populations treated by different models of care. It is likely that true differences do exist for patient sub-groups on specific care pathways in particular contexts but the signal of these differences is hidden by system-level analyses for whole populations. Our studies provide useful information about aggregate outcomes and trends over time but more focused studies are required to provide evidence to policy makers about the impact of system changes such as hospital reconfiguration. These studies should focus on processes, such as compliance with new hospital referral guidelines, as well as outcomes[13]. It is also possible that the impact of interventions in complex systems may not be apparent in the timeframes we used, and that longer-term studies are required[13].

Fourth, each indicator is targeted at a heterogeneous group of patients and is a composite of different outcomes. This can hide valuable information about patient sub-groups and specific outcomes. The case-fatality study, for example, combines outcomes for 16 different conditions[10]. The patient experience study combined multiple questions into composite scale scores[12]. Both of these found no signal of outcome differences between models of urgent and emergency care which may indicate that a finer grain of analysis is required when using performance indicators for urgent and emergency care systems.

Fifth, routinely collected secondary data can rarely be used to calculate the indicators presented in this paper and original data collection is often required. Routine data could only be used for two of the 17 indicators (case-fatality ratios for serious emergency conditions and admission rates for ambulatory conditions). The study of patient experience was original because no such data was available at the time it was conducted. Ireland has since introduced a national patient experience survey which is conducted annually[14]. However, this still focuses on individual emergency departments and would not, therefore, be suitable for an analysis of the impact of whole system changes on patient experience. Time-based indicators are almost completely absent from the health information systems in public hospitals in Ireland.

Finally, one of our studies found that the indicator 'emergency admission rates for conditions whose exacerbations could be managed out of hospital or in ED's without admission to an inpatient bed' was no more sensitive to variations in care settings than total emergency admission rates. This underlines the importance of a critical approach to performance indicators, which may not be fit for purpose despite seeming to have obvious face validity.

Ethical statement
Ethical approval for the study was granted by the Clinical Research Ethics Committee of the Cork Teaching Hospitals [ECM 4 (q) 02/07/13]. The process of participants proceeding to the survey and completing it constituted consent.

Data availability
The data is available on Open Science Framework: http://doi.org/10.17605/OSF.IO/3CW6F[8]

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Supplementary material
Supplementary File 1: PRISMA checklist.
Click here to access the data.

Supplementary File 2: All novel indicators identified in the updated systematic review combined with those identified in the original systematic review.
Click here to access the data.

Supplementary File 3: Definitions of urgent care-sensitive and emergency conditions.
Click here to access the data.

Supplementary File 4: Median, range and mean of consensus development group ratings for all indicators following consensus development group meeting on a 9 point Likert scale. Higher scores represent a higher level of agreement that the indicator is a useful measure of system performance.
Click here to access the data.

References

1. Carson D, Clay H, Stern R: **Breaking the mould without breaking the system.** Primary Care Foundation. 2011.
    Reference Source

2. Ham C, Smith J, Temple J: **Hubs, spokes and policy cycles: a paper for the King's Fund London Commission.** King's Fund. 1998.

3. Nicholl J, Coleman P, Jenkins J, *et al.*: **MCRU Programme 2006–2010. The emergency and urgent care system.** Final Report. University of Sheffield. 2011.
    Reference Source

4. 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.
    PubMed Abstract | **Publisher Full Text**

5. Nicholl J, Coleman P, Knowles E, *et al.*: **MCRU Programme 2006–2010. The emergency and urgent care system.** Final Interim Report of Phase 2006–2008. University of Sheffield. 2009.
    Reference Source

6. Zalstein S, Danne P, Taylor D, *et al.*: **The Victorian major trauma transfer study.** *Injury.* 2010; **41**(1): 102–9.
    PubMed Abstract | **Publisher Full Text**

7. Paoloni R, Fowler D: **Total access block time: a comprehensive and intuitive way to measure the total effect of access block on the emergency department.** *Emerg Med Australas.* 2008; **20**(1): 16–22.
    PubMed Abstract | **Publisher Full Text**

8. Murphy A, Wakai A, Walsh C, *et al.*: **Development of key performance indicators for prehospital emergency care.** *Emerg Med J.* 2016; **33**(4): 286–92.
    PubMed Abstract | **Publisher Full Text**

9. Khalifa M, Zabani I: **Developing Emergency Room Key Performance Indicators: What to Measure and Why Should We Measure It?** *Stud Health Technol Inform.* 2016; **226**: 179–82.
    PubMed Abstract | **Publisher Full Text**

10. Lynch B, Fitzgerald AP, Corcoran P, *et al.*: **Case fatality ratios for serious emergency conditions in the Republic of Ireland: a longitudinal investigation of trends over the period 2002-2014 using joinpoint analysis.** *BMC Health Serv Res.* 2018; **18**(1): 474.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Lynch B, Fitzgerald AP, Corcoran P, *et al.*: **Drivers of potentially avoidable emergency admissions in Ireland: an ecological analysis.** *BMJ Qual Saf.* 2018; pii: bmjqs-2018-008002.
    PubMed Abstract | **Publisher Full Text**

12. Foley C, Droog E, Boyce M, *et al.*: **Patient experience of different regional models of urgent and emergency care: a cross-sectional survey study.** *BMJ Open.* 2017; **7**(3): e013339.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Moore GF, Evans RE, Hawkins J, *et al.*: **From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation.** *Evaluation (Lond).* 2019; **25**(1): 23–45.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

14. Health Information and Quality Authority: **National Patient Experience Survey. Findings of the 2018 Inpatient Survey.** 2018.
    Reference Source

Open Peer Review

Current Peer Review Status: ✔ ❓ ❓ ✔

Version 2

Reviewer Report 26 June 2019

✔

Anna Cantrell 🆔

School for Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK

The abstract contains all of the relevant information. The results section description of the literature flow chart does not agree with the flow chart of the literature review, more detail below.

The introduction section provides a clear introduction to the topic.

The methods for the systematic review are clear. It might have been useful for the authors to search another database in addition to Medline, perhaps Embase, Cochrane Library, CINAHL or HMIC but the decision to just search Medline might have been due to time, resource and access constraints. It would have been useful to include the PubMed search strategy in format that would be reproducible, this could be as supplementary material if there is not space in the article. The sources contacted for grey literature are appropriate and the search for and inclusion of grey literature is a strength of the review. I am not concerned by the dates of the search as the articles retrieved were used to develop a list of novel indicators for the consensus development exercise which has now occurred and it is discussed as a limitation that the systematic review only covers up to 2014 and more recent literature could have been missed.

The inclusion of the PRISMA checklist is useful but I don't think that the information about the search strategy is sufficient to enable it to be repeated. The PRISMA checklist states "Present full electronic search strategy for at least one database", this article only includes the statement by combining variants of the text terms 'emergency' and 'indicator' using the AND operator.

The methods of the consensus development exercise are again clear. There is detailed information about the broad range of experts involved.

The results section first describes the results of the literature search and is reasonably clear but hard to follow in conjunction with the flow chart. The flow chart for the literature review is difficult to follow and hard to correlate with the results section in the abstract. For example, the chart has the 2339 articles from the literature search then the results sections mentions further searches and then 60 items were identified for full-text review but it is 47 in the flow chart. Completing The PRISMA 2009 Flow Diagram from here

would have been more appropriate and then hopefully clearer. The results of the consensus development exercise are clear.

The study limitation section is appropriate, all the limitation mention are relevant and described sufficiently.

The implications section is well organised starting by discussing three studies that have been been conducted using the results of this article then the implications for research.

Overall, a good interesting paper.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Information science, Systematic reviewing.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 11 June 2019

https://doi.org/10.21956/hrbopenres.13970.r26625

?   **Maaret Castrèn**
    Department of Emergency Medicine and Services, Helsinki University Hospital and Helsinki University, Helsinki, Finland

After reading the reviewers reports I still have some concerns. It would have benefited the article if the search would have been done for more new articles, 2014 seems like a long time ago. And even if you explain this later, the reader sees 2014 as the deadline.

You state that you exclude individual components but still you have some in the indicators so this remains unclear for the reader. You never really explain what happened to the 6 grey literature other documents. Figure 1 should be written in a different format, now you have it like the flow would be from the 2339 potentially relevant articles and of course this is not the case so you need to make three different flows in this figure.

It is not clear in Table 1 what EUCS is so write it out and also SIREN work package is unclear for the reader. I don't understand why you include inter-hospital transfers, most of them being not acute or emergency medicine.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 15 March 2019

✔   **Robbert Huijsman** 🆔
Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

I propose to accept the paper, although it is a pity that the authors did not follow the suggestion of both reviewers to extend the literature search which now stops in 2014 (5 years ago!).

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 01 May 2018

https://doi.org/10.21956/hrbopenres.13865.r26092

? **Ellen Weber**

Department of Emergency Medicine, University of California, San Francisco, San Francisco, CA, USA

Thank you for the opportunity to review this paper. It is nicely written and succinct. The research question is important and novel; there have several studies in the field of prehospital care and emergency medicine looking at performance indicators but they usually focus on service in the system. The idea of finding indicators that reflect the performance of the whole system is important and could lead to alignment of incentives that currently do not exist, in both Ireland and beyond. Overall the study is well-done. I have a few concerns about a few of the indicators representing (whole system) see later, but overall, the indicators seem to represent the goals the authors set out to achieve and many have not been

emphasized previously.

I have several suggestions to strengthen the paper. First, the systematic review only goes to 2014. Since then there have been at least two other Delphi studies on indicators in emergency medicine[1,2]. This is a limitation if the only indicators assessed were those found in these reviews (i.e. none submitted by the experts). Although I understand that the study rests on what was available at the time, I would improve the strength of the paper if the authors could do an updated literature review and see if there are any other indicators that might have been included had these newer articles been in the systematic review. Additionally, it would helpful if the authors could summarise the newly added articles in a list in an appendix with articles, dates, authors, location, service(s) involved, etc.

2). The methods by which the articles and the performance indicators were initially chosen is not well described. How many individuals were involved in the article selection, and how was the process of extracting the new indicators from the articles done. Was there any cross-checking among the investigators? I would imagine that there would be some areas of judgment here as to whether an indicator was new or differently stated, as well as whether it applied to the whole system. Was there any type of inter-rater reliability assessed?

3) The discussion should be expanded to include comparisons of this work with prior processes within emergency medicine to identify key performance indicators. You have a brief mention of the UK exercise, although this comes a bit too early in the discussion, and without much detail. (You also say that 6 were from your systematic review but you originally identified only 3 new ones compared with the prior review..) It would help to point out what indicators were novel, which part of the systematic review they originated in, and why you think your indicators were different. It would also help to a discussion of the pragmatic implications of this work – how have such performance indicators been use to evaluate systems of care and how do the authors imagine theirs being used in the context of Ireland's reconfiguration.

A few other points that need clarity:

1. The description of the article search and selection does not readily match the flow diagram. EG the flow diagram doesn't show the 150 articles that had abstracts read. And then the 47 articles that were reviewed.

2. How was the grey literature identified?

3. Were all experts from Ireland?

4. There were 17 experts in the total group, and you say that 13/14 attended the meeting. What happened to the other 3?

5. Could you explain the procedure at the meeting a bit more – did each of the 13 participants speak about each of the indicators. (Sounds like a very long day!) What is meant by a "nominal" group technique.

6. The lists of indicators both in the main text and the appendix are confusing. In Table 1 you have 17 indicators which you said had at least a median score of 7 in the meeting and they were then ranked in an on-line process. It is not clear how those with a median of 2 and 3 are included. I am also not sure the reader needs to see the "mean rank" but rather the final ranking order based on all responses. Even if you keep those ranks in, it would help to organize the table by ranking.

Additionally, you could include a notation for which are process and outcome measures, and which (if any) were from the updated systematic review. In the appendix, its not clear you need to have two separate lists of all measures. Could you combine them, and present them according to type (process, structure, outcome).

7. I wonder if the authors would also care to discuss some of the indicators which don't seem to this reviewer to represent "whole system." Performance. E.g. Call to ambulance to time on scene. While this is described as being useful to measure variation, it is still very "service" specific – regarding the prehospital system. Similar question for "Time from onset of serious emergency condition to arrival at the receiving hospital".

8. Abstract – The methods of the abstract should be expanded to explain a bit more about the updating of a prior review, and more about the consensus exercise, including the variety of experts. Provide a brief explanation of "nominal".

Minor editorial issues:

Page 3, third paragraph – the last sentence – "This variation represents…." Is probably unnecessary and takes the introduction a bit off topic.

Page 3, first sentence of methods – Please add the publication date of the first systematic review to the text so readers don't have to search your references.

Page 4, second paragraph – should be indicators

Page, 4 end of third paragraph, the" This review led to three articles being identified for inclusion in the review" - might change to: This process led to …

References
1. Murphy A, Wakai A, Walsh C, Cummins F, O'Sullivan R: Development of key performance indicators for prehospital emergency care.*Emerg Med J*. 2016; **33** (4): 286-92 PubMed Abstract | Publisher Full Text

2. Khalifa M, Zabani I: Developing Emergency Room Key Performance Indicators: What to Measure and Why Should We Measure It?. 179-182 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Health policy, ED operations, information technology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 23 April 2018

https://doi.org/10.21956/hrbopenres.13865.r26158

?   **Robbert Huijsman** (iD)

Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

Please add in subtitle that it's an <u>update of a</u> systematic review. How did you verify your results with those of the authors of the previous review (and for instance check your decision to exclude 7 papers from that review). Your update is already slightly outdated, because you performed it in August 2014. Although you were quick with the consensus group meeting in October 2014, directly after your review update, now it is more than 3 years later. Please reflect on this in your Discussion, especially as the acute field is evolving rapidly. Also discuss the generalisability of your findings for Ireland to other countries with different systems of emergency and acute care.

The number of experts in your consensus group goes down in two steps from 17 (bottom line on page 2), to 14 and then 13. Please explain in more detail the reasons for dropout in two steps, and possible impact on your findings.

In the Results, your text about excluded articles does not parallel the semi-final box in Figure 1, please synchronize. The way of scoring is not totally clear to me. How did the ranking actually take place and what then is the meaning of median scores and the differing score ranges per indicator? Your threshold of 7 has a different meaning if indicators vary in their range, so the order in the table is not that self evident. Experts ranked the indicators twice, it seems. In the first round you selected indicators with a median higher than 7, but how then can Table 1 about the second round of ranking include two indicators with a median of (much) lower than 7? You added * and ** to some entries in Table 1, but their is no explanation below the table.

Your Discussion is rather short and poor, please elaborate more on the content of your findings, in combination with the existing literature and professional guidelines, not just restrict yourselves to a (too short) discussion on the methods. For instance, your recommendations on future steps about

measurement are rather superficial. Perhaps reflect on a Registry, a quality system with feedback and improvement cycles, or other ways to incorporate performance indicators in the real work of professionals and managers in emergency care.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.