BMJ Global Health

Use of interrupted time series methods in the evaluation of health system quality improvement interventions: a methodological systematic review

Celestin Hategeka ,[1,2] Hinda Ruton,[2,3] Mohammad Karamouzian,[4,5] Larry D Lynd,[6,7] Michael R Law[2]

To cite: Hategeka C, Ruton H, Karamouzian M, *et al*. Use of interrupted time series methods in the evaluation of health system quality improvement interventions: a methodological systematic review. *BMJ Global Health* 2020;**5**:e003567. doi:10.1136/bmjgh-2020-003567

Handling editor Valery Ridde

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjgh-2020-003567).

Received 28 July 2020
Revised 7 September 2020
Accepted 10 September 2020

For numbered affiliations see end of article.

Correspondence to
Dr Celestin Hategeka;
chategeka@hsph.harvard.edu

ABSTRACT

Background When randomisation is not possible, interrupted time series (ITS) design has increasingly been advocated as a more robust design to evaluating health system quality improvement (QI) interventions given its ability to control for common biases in healthcare QI. However, there is a potential risk of producing misleading results when this rather robust design is not used appropriately. We performed a methodological systematic review of the literature to investigate the extent to which the use of ITS has followed best practice standards and recommendations in the evaluation of QI interventions.

Methods We searched multiple databases from inception to June 2018 to identify QI intervention studies that were evaluated using ITS. There was no restriction on date, language and participants. Data were synthesised narratively using appropriate descriptive statistics. The risk of bias for ITS studies was assessed using the Cochrane Effective Practice and Organisation of Care standard criteria. The systematic review protocol was registered in PROSPERO (registration number: CRD42018094427).

Results Of 4061 potential studies and 2028 unique records screened for inclusion, 120 eligible studies assessed eight QI strategies and were from 25 countries. Most studies were published since 2010 (86.7%), reported data using monthly interval (71.4%), used ITS without a control (81%) and modelled data using segmented regression (62.5%). Autocorrelation was considered in 55% of studies, seasonality in 20.8% and non-stationarity in 8.3%. Only 49.2% of studies specified the ITS impact model. The risk of bias was high or very high in 72.5% of included studies and did not change significantly over time.

Conclusions The use of ITS in the evaluation of health system QI interventions has increased considerably over the past decade. However, variations in methodological considerations and reporting of ITS in QI remain a concern, warranting a need to develop and reinforce formal reporting guidelines to improve its application in the evaluation of health system QI interventions.

Key questions

What is already known?

- Interrupted time series (ITS) is one of the strongest quasi-experimental designs.
- ITS has increasingly been advocated for use in the evaluation of health system quality improvement (QI) interventions when randomisation is not possible.

What are the new findings?

- Our findings showed variations in methodological considerations and reporting of ITS in QI—for example, only about half of studies specified the ITS impact model and considered autocorrelation in the ITS analyses.
- The risk of bias was high or very high in about three-quarters of included studies and did not appear to have changed significantly over time.

What do the new findings imply?

- While use of ITS in the evaluation of health system QI interventions has increased considerably over the past decade, the quality of ITS studies still shows room for improvement.
- Developing and/or reinforcing formal reporting guidelines could improve application of ITS in the evaluation of health system QI interventions.

INTRODUCTION

Improving health system quality is vital to ensure that utilisation of healthcare maximises population health outcomes, economic benefit and people's confidence in health systems.[1–3] A key strategy to improving the quality of care is the use of context-appropriate health system quality improvement (QI) interventions. QI interventions employ different approaches targeting patients or users (eg, education of patients and promotion of self-management), healthcare providers (eg, clinical education and reminders) and health systems (eg, organisational change and electronic patient registry) to bolster quality of care.[4 5] While evaluating these interventions is key to understanding their impact, many are designed and

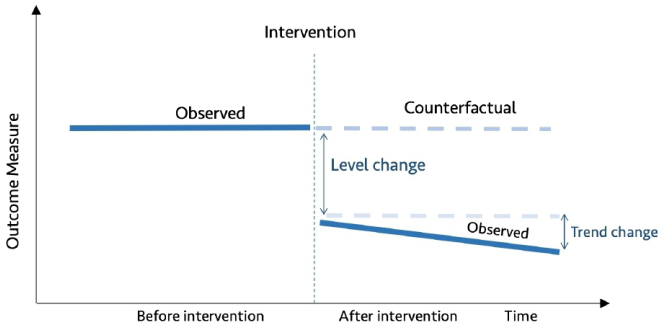

Figure 1 Diagrammatic representation of single interrupted time series.

implemented without predetermined evaluation plans.[6–9] In such cases, rigorous quasi-experimental designs can be employed.[10–13]

Concerns persist, however, about the potential biases in some weaker observational study designs.[10 11] As a result, interrupted time series (ITS) analysis has increasingly been advocated as one of the more robust observational designs as it can control for the secular trends present in many health system outcomes.[7 8 14] Additionally, ITS can easily be interpreted by lay audiences, especially when graphical displays of the results are employed. Similarly, the Cochrane Effective Practice and Organisation of Care (EPOC) recommends ITS along with experiments and controlled before and after designs for inclusion in EPOC reviews.[15] There remains a risk, however, of producing misleading or incorrect results when this rather robust design is not used appropriately.[16–18]

While use of ITS in health research has increased substantially over the last few decades (online supplemental appendix figure 1), previous systematic reviews have shown heterogeneity in the reporting and methodological concerns in the application of ITS.[17 19–21] However, none of these reviews have focused on use of ITS in the evaluation of health system QI interventions.[17 19 20] Another review examined characteristics and reporting of ITS in healthcare more broadly.[21] However, this review was restricted to just 1 year (ie, 2015) and one database and excluded controlled ITS studies.[21] Similarly, they did not assess the risk of bias of included studies. Therefore, the extent to which the use of ITS has followed best practice standards and recommendations in the evaluation of health system QI interventions remains unclear.

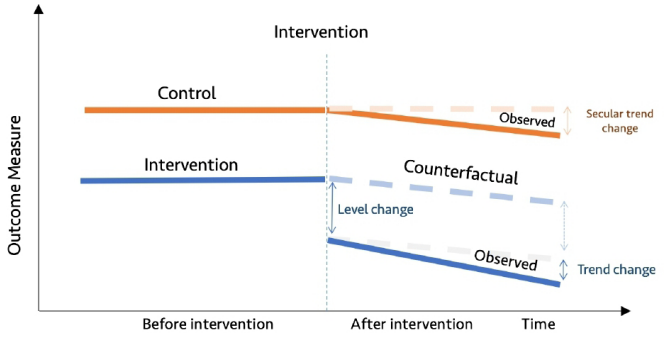

Figure 2 Diagrammatic representation of controlled interrupted time series.

Methodological considerations with interrupted time series

The ITS design relies on data collected at multiple intervals over time (ie, time series data) before and after an intervention to establish a causal relationship between an intervention (eg, QI) and an outcome of interest (eg, health outcomes).[22] The Cochrane EPOC recommends ITS studies with at least three data points before and after an intervention for inclusion in its reviews.[15] ITS can be either retrospective or prospective, and can further be classified into two types based on whether one group was used (single ITS) or a control group was added (controlled ITS).[10 22] Single ITS (SITS) compares longitudinal changes before and after the intervention in the exposed group, whereas controlled ITS (CITS) compares longitudinal changes before and after the intervention between an exposed group and a control group (figures 1 and 2).[7 18 22–25]

Single ITS assumes that level and trend in a given outcome measure in the group exposed to the intervention would have remained the same absent the intervention (figure 1). In contrast, the counterfactual assumption for controlled ITS is that the level and trend in the group exposed to the intervention would have changed in the same way as was observed in the control group (figure 2).[7 18 22–25] It is recommended to specify anticipated changes in the level or trend in outcome measures a priori, for example, based on expert knowledge.[17 26 27] Graphical figures of results of ITS studies are vital as they allow readers to see pre-existing levels and trends, the implementation of the intervention and the impact of the intervention.[18 20 28]

CITS is arguably a stronger quasi-experimental design as it has two controls (baseline trend in the intervention group and control group), allowing both within-group and between-group comparisons.[10] This enhances its capacity to control for potential threats to internal validity such as history (co-intervention), maturation, instrumentation, regression to the mean and attrition.[10 18 22 24 29] However, SITS may be more vulnerable to these threats to internal validity. For example, Baicker and Svoronos recently reanalysed data from the Oregon Health Insurance experiment using SITS and found results that were not consistent with the randomised trial findings.[30] Therefore, it is generally recommended to add a control group to ITS wherever possible.[10 22] Bernal and colleagues' study reviewed various types of controls that can be added to strengthen ITS design, including location and characteristic-based controls and control outcome.[24] When QI interventions are multisite and implemented at different times, this needs to be taken into account when assigning index time to a control group using techniques such as matching and random assignment of index time to controls.

Statistical analysis of interrupted time series

While there are many statistical approaches used for estimating effects of interventions in ITS studies, segmented regression analysis of ITS is the most commonly used, followed by autoregressive integrated moving average (ARIMA) models.[16 20 21] Details on segmented regression analysis can be found in the seminal paper of Wagner and colleagues.[18] While most ITS studies use segmented regression analysis on time series data aggregated within periods of time, analysis on individual-level data is feasible using mixed effects models. Saeed and colleagues discuss how to combine generalised mixed effects models with segmented regression to evaluate the effectiveness of healthcare interventions.[31] Similarly, Ewusie suggests incorporating patient variability and sample size as weights in a weighted segmented regression model of ITS to account for potential heterogeneity introduced by population variability within and/or between sites.[16] Power in ITS depends on several factors, including the number and distribution of data points before and after the intervention, the stability of data points, strength of effect, autocorrelation and seasonality.[26 32] Missing observations are common in time series data of health service utilisation and quality of care and should be dealt with using appropriate techniques such as multiple imputation and maximum likelihood estimation.[25 33–35]

As successive time series observations may be correlated, it is recommended to assess for autocorrelation using appropriate modelling techniques.[18] There are several tests for assessing autocorrelation including Durbin-Watson tests and visual plots of the autocorrelation function and partial autocorrelation function.[18 36] Moreover, seasonality is common in time series data and should be adjusted for as appropriate. There are several techniques for adjusting for seasonality in time series analyses, including deseasonalising time series data prior to fitting statistical models, adding dummy calendar variable (eg, calendar month), using Fourier terms (pairs of sine and cosine functions) and splines.[26 37] When using monthly time series data, at least 12 data points before and after the intervention are recommended to be able to adjust for seasonality meaningfully.[18] Additionally, non-stationarity—a time series exhibits an underlying trend that is not related to the intervention—should be assessed and, where significant, stationarised as appropriate. Non-stationarity can be assessed using the Augmented Dickey-Fuller test.[38] When autocorrelation, seasonality and/or non-stationarity are present in time series data and are not adjusted for, this can result in biased results.

Despite the rich literature on how to approach ITS studies from a methodological and statistical standpoint, we do not have good information on the degree to which evaluations of healthcare QI interventions meet these criteria. Thus, we systematically reviewed the literature to investigate the extent to which these evaluations have followed best practice standards.

METHODS

This methodological systematic review was conducted according to a prespecified study protocol registered in PROSPERO (registration number: CRD42018094427) on 11 June 2018.[39]

Search strategy

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (online supplemental appendix table 1),[40] we searched for healthcare QI intervention studies that were evaluated using ITS and were published in peer-reviewed journals from inception to June 2018. We also searched the grey literature to capture any relevant studies. The search strategy is summarised in table 1 and full details are provided in online supplemental appendix table 2.

Inclusion and exclusion criteria

The population, interventions, comparisons, outcomes and study designs considered for review are listed in table 2. Guided by the Shojania *et al*'s taxonomy for QI

Table 1 Summary of the search strategy

Search concepts*	Interrupted time series analysis (eg, interrupted time series analysis OR ITS studies OR interrupted time series OR time series OR trend analysis OR segmented regression OR Piecewise regression OR broken-stick regression) AND healthcare QI interventions (eg, quality improvement OR healthcare quality OR quality control OR quality assurance OR clinical audit).
Databases	MEDLINE, EMBASE, CINAHL, Web of Science, Global Health, Google Scholar, Africa-Wide, Latin American and Caribbean Health Sciences Literature (LILACS), Index Medicus for the South-east Asian Region (IMSEAR), Index Medicus for the Eastern Mediterranean Region (IMEMR), and Western Pacific Rim Region Index Medicus (WPRIM).
Other search strategies	Complementing electronic databases, hand searches of the bibliographies of relevant published works and previous reviews, relevant conference proceedings (eg, International Forum on Quality and Safety in Healthcare, Institute of Healthcare Improvement) were also performed.
Restrictions	No restrictions applied on date of publication, language of study, participants, or the type of QI outcome.

*Search terms were combined using appropriate Boolean operators and included subject heading terms and/or key words for two key themes (interrupted time series analysis and healthcare quality improvement) and were adjusted to fit each database requirements.
ITS, interrupted time series; QI, quality improvement.

Table 2 Population, interventions, comparisons, outcomes and study designs for study inclusion

Criteria	Definition
Population	All types of patients/consumers and healthcare professionals/providers are eligible. All levels of healthcare delivery are eligible.
Intervention	Interventions to improve quality of healthcare[4 5]: ► Provider reminder systems; ► Facilitated relay of clinical data to providers; ► Audit and feedback; ► Provider education; ► Patient education; ► Patient reminder systems; ► Promotion of self-management; ► Organisational change; and ► Financial incentives, regulation and policy
Comparison	Not relevant given this review is not focused on any particular comparator.
Outcomes	Not relevant given this is a methodological review.
Study designs	Interrupted time series

strategies, we identified interventions targeting health systems, health providers and/or patients for inclusion.[4 5] Studies that did not have at least three preintervention and postintervention data points were excluded, consistent with the Cochrane EPOC recommendation.[15] Similarly, given that this review focused on empirical applications of ITS, systematic reviews/meta-analyses, letters, commentaries, methodological contributions, study protocols and conference abstracts were excluded.

Data extraction and analysis

The titles and abstracts of unique results from all databases and grey literature were reviewed independently by two researchers (CH and HR) for potential inclusion. The full texts of studies retained at the title and abstract screening stage were retrieved and independently assessed for inclusion by two researchers (CH and HR). Any discrepancies were resolved through discussion and consensus, with a third researcher (MRL) consulted wherever necessary. Data extraction on each included study was conducted using a data extraction tool, developed and piloted a priori drawing on a checklist on the methodological and reporting recommendations for ITS studies adapted from the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (online supplemental appendix table 3).[20] Data elements included study characteristics (eg, publication year, country), intervention characteristics (eg, QI strategies, single/multiple interventions), methodological considerations (eg, ITS model, autocorrelation, use of control). We synthesised extracted data narratively using descriptive statistics and following the review protocol registered in PROSPERO. Data extraction was done by one researcher (CH); however, a validation was done by a second researcher (HR) who verified a random subsample consisting of 10% of eligible studies. The agreement between the two researchers (CH and HR) was very high (>95%). Additionally, where necessary, CH consulted with the researcher team members during data extraction. Lastly, assessment of the risk of bias of included studies was performed by two researchers (CH and HR).

Assessment of risk of bias

The risk of bias assessment was guided by the Cochrane Effective Practice and Organisation of Care (EPOC) standard criteria for assessing risk of bias for ITS studies.[17 27 41] Specifically, we assessed the following EPOC standard criteria:

1. Was the intervention independent of other changes?
2. Was the shape of the intervention effect prespecified?
3. Was the intervention unlikely to affect data collection?
4. Was the primary outcome measured objectively?
5. Were incomplete outcome data adequately addressed?
6. Was the study free of selective outcome reporting?
7. Was the study analysed appropriately using interrupted time series techniques?

Criteria 1, 3 and 4 assess the threat of history, instrumentation and testing, respectively. Criterion 2 assesses whether the ITS impact model was specified a priori, while criterion 5 assesses whether missing data were dealt with appropriately. Criterion 6 assesses whether all relevant outcomes that were part of the study objectives were reported. Lastly, criterion 7 assesses whether data were analysed appropriately such as using appropriate models (eg, segmented regression and ARIMA) and considering key methodological recommendations (eg, autocorrelation and seasonality).

Online supplemental appendix table 4 describes how we performed risk of bias scoring among included studies. Each criterion scored '0' if low risk and '1' otherwise. For each study, we created an aggregate score by combining scores across the seven criteria. Subsequently, the study-level risk of bias aggregate score was categorised as low (score=0), moderate (score=1 or 2), high (score=3 or 4) and very high (score >4). Trends in risk of bias over time were also assessed. The time for each included study was defined as the year of the publication. Time trend for risk of bias was assessed by fitting a logistic regression model comparing the proportion of studies with a low or moderate risk of bias to the studies with high or very high risk of bias as has previously been done.[42] As the literature search end date was June 2018, studies published in 2018 that were captured by our search strategy may not have been representative of all studies published in 2018. As such, for time trend analyses, studies published in 2018 were excluded. All statistical analyses were conducted using R version 4.0.2

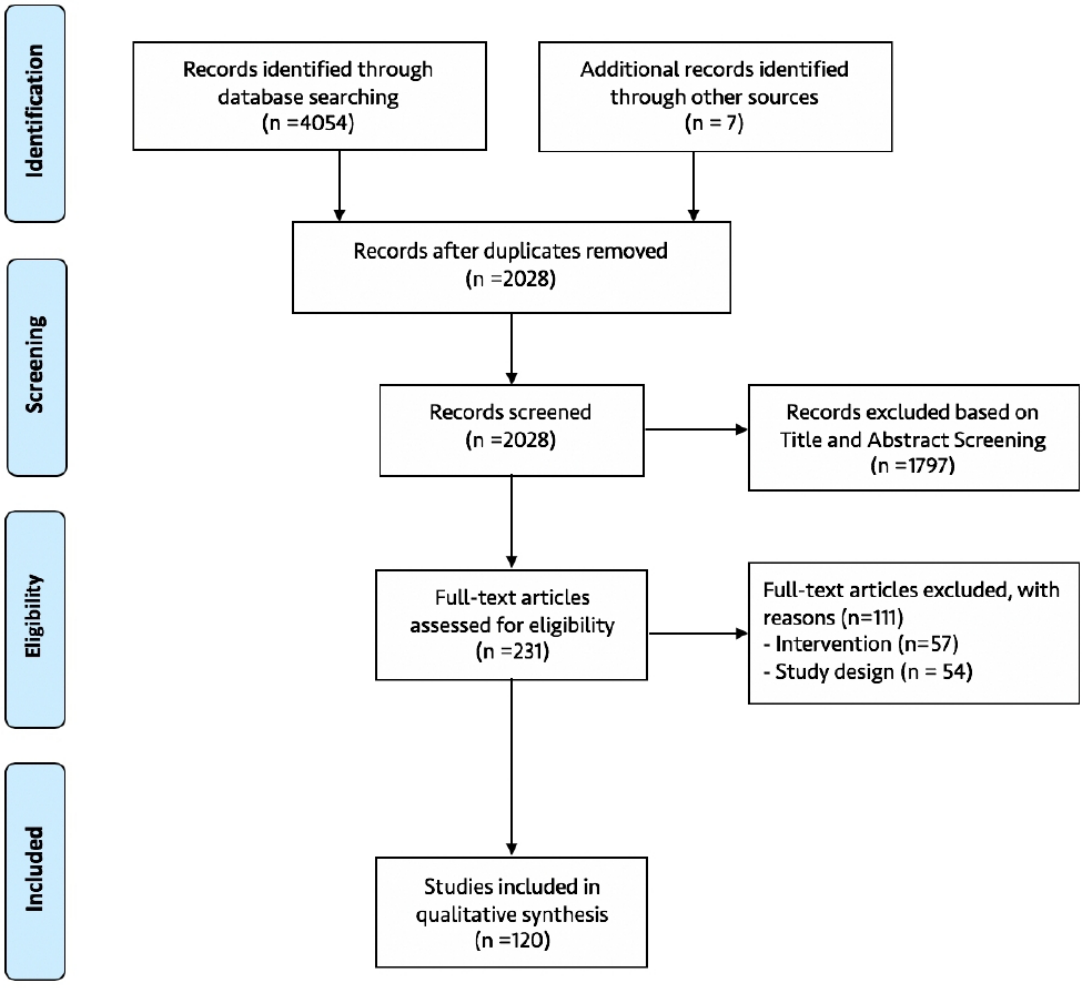

Figure 3 Flow diagram of the selection of included studies.

Patient and public involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

RESULTS

The search strategy identified 4061 results, of which 2028 unique eligible records were screened for inclusion. Abstract and full-text screening identified 120 studies that met the inclusion criteria.[43–162] A summary of this process is presented in the PRISMA flow diagram in figure 3.

Characteristics of included studies

The 120 studies represented a diversity of QI strategies from different geographical settings. Included studies were from 25 countries, with 106 (88.3%) from high-income countries and 14 (11.7%) from low- and middle - income countries (LMICs) as defined by the 2019 World Bank Classification (online supplemental appendix figure 2). The number of studies published over time increased considerably, with 104 (86.7%) studies published in 2010 or later (online supplemental appendix figure 3). The included studies assessed eight QI strategies, with provider education (39.2%), audit and feedback (30.0%), organisational change (24.2%) and provider reminder systems (19.2%) being the most reported QI strategies (figure 4). Nearly half of studies were multisite (44.2%).

Reporting and methodological considerations

Interrupted time series methodological consideration of included studies

All included studies reported the use of an ITS design in their titles and/or abstracts and provided clear timing related to the QI implementation (table 3). Nearly all studies provided a background/rationale for using ITS (98.3%) and provided a description of the intervention (96.7%). A fifth of studies reported on evaluation of multiple phases of QI interventions, of which 76.0% adjusted for multiple phase of QI interventions in their analyses. All included studies reported on the study population, inclusion criteria and outcomes. Nearly all studies (99.2%) used data collected at regular intervals and reported data sources, of which 57% were administrative

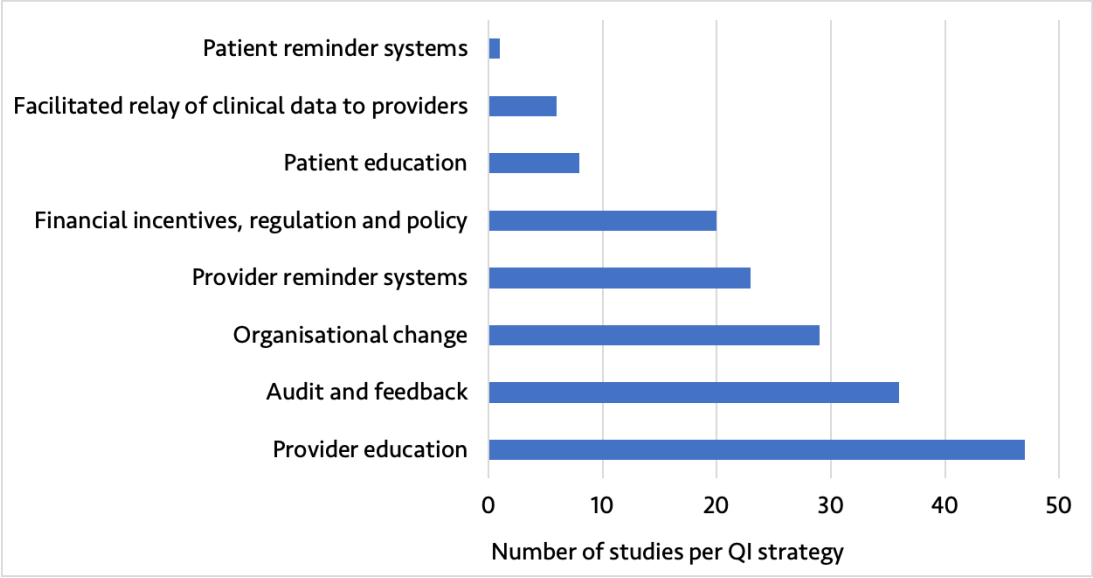

Figure 4 Type of quality improvement (QI) strategy reported in the included studies.

data including Health Management Information System data, 35% were patient records and 8% were survey data.

Most studies (71.4%) used monthly data points, with a median preintervention and postintervention period of 18 months (range, 3–120) and 20 months (range 4–90), respectively (table 3). Less than half (42.5%) of studies provided a rationale for the number and spacing of data points described. Just over half reported the unit of analysis clearly (42.5% used data aggregated at the study site level and 9.2% used disaggregated data). Most studies used ITS without a control (81%) and modelled data using segmented regression (62.5%), while 1 in 10 (11.6%) did not report what model was used. For studies that included a control group, 59.1% of studies used a location-based control and about half (54.5%) modelled intervention and control group in a separate ITS model.

About half (55%) of studies reported checking/adjusting for autocorrelation in their ITS models, with the Durbin-Watson test statistic being the test frequently used to check for autocorrelation (36.4%). A fifth (20.8%) of studies reported checking and adjusting for seasonality. Similarly, only 10 (8.3%) studies reported checking/adjusting for non-stationarity in their ITS models, with the augmented Dickey-Fuller test being the most frequently used (70%). Stratified analysis showed that these items were more likely to be reported with ARIMA models compared with segmented regression (online supplemental appendix table 5). Only about half (49.2%) of studies specified the ITS impact model. Over a tenth (12.5%) of studies considered sensitivity analyses.

Reporting of interrupted time series study results and interpretation

Most studies reported characteristics of study sample (70%) and reported all outcomes examined over the study period (87.5%) (table 4). Similarly, most studies reported the impact of QI using level and trend changes (89.2%), reported CI or SE (81.7%), and reported graphical figures displaying results (92.5%). Of the studies that included graphical figures, 57.6% showed observed trends and 12.6% showed the counterfactual. Further, 29.2% of studies reported absolute and/or relative changes and their significance, for example, clinical and policy. Most studies reported key results and discussed potential limitations of the study. Less than half of studies explicitly discussed potential history threats such as co-interventions. Similarly, 42.1% of studies that mentioned potential limitations also discussed the direction or magnitude of these potential bias(es). About a quarter of studies discussed the stability of outcome data over time.

Risk of biases of included studies

Risk of bias was low for 3 (2.5%) studies, moderate for 30 (25.0%), high for 64 (53.3%) and very high for 23 (19.2%). The proportion of studies with a low or moderate risk of bias did not significantly change over time (OR 1.15, 95% CI 0.96 to 1.40, p=0.14) (figure 5). Figure 6 shows the risk of bias across the seven EPOC standard criteria for assessing bias in ITS studies. Details on risk of bias assessment for each included study are provided elsewhere.[6]

DISCUSSION

Over the past few decades, ITS has been an increasingly popular method to evaluate QI interventions in healthcare. We investigated whether the use of ITS in such evaluations has followed best practice standards and recommendations. Our findings showed variations in methodological considerations and reporting of ITS in QI. In particular, we found that most studies identified used ITS without a control and only about half of studies specified the ITS impact model and considered autocorrelation in

Table 3 ITS methodological considerations of included studies

Characteristics	n	%
ITS design reported in the title and/or abstract	120	100
Background/rationale reported	118	98.3
Study objectives reported	120	100
Description of QI intervention	116	96.6
Start (and end) of QI intervention reported	120	100
Multiple QI interventions	25	20.8
► 2 interventions	22	88
► 3 interventions	1	4
► 4 interventions	2	8
Adjusted multiple interventions in the analyses	19	76
Study setting reported	120	100
Multisite	53	44.2
Study period reported	120	100
Study population reported	120	100
Cohort definition reported	120	100
Inclusion criteria reported	120	100
Data sources reported	119	99.2
Data completeness and validity reported	14	11.6
Time point intervals, monthly	85	71.4
Data collected regularly (regular interval)	119	99.2
Time points clearly reported	119	99.2
Rationale for the number and spacing of data points described	51	42.5
Outcome measure(s) reported	120	100
Format of outcome(s) reported	120	100
Unit of analysis		
► Aggregated	51	42.5
► Individual	11	9.2
► Not reported	58	48.3
ITS models		
► Segmented regression	75	62.5
► ARIMA	19	15.8
► Other models (eg, GEE, linear regression, mixed effect model, spline regression, poison regression, Prais-Winsten regression, logistic regression)	17	14.2
► Not reported	14	11.6
Autocorrelation		
► Checked/adjusted as appropriate	66	55
Test(s) used to check for autocorrelation		
► Durbin-Watson test statistic	24	36.4
► Other tests (eg, ACF and PACF, Ljung-Box χ^2 test, residual plots and Breusch-Godfrey test, Breusch-Godfrey test, Cumby-Huizinga test)	10	15.2
► Not reported	33	50
Seasonality		
► Checked/adjusted	25	20.8
Non-stationarity		
► Checked/adjusted	10	8.3
Test(s) used to check		
► Augmented Dickey-Fuller test	7	70
► Not reported	2	20
Control group used	22	18.3
Type of control, location-based control	13	59.1
Analyses of controlled ITS		
► Combined	5	22.7
► Separated	12	54.5
► Difference	2	9.1
► Not reported	3	13.6
Specify ITS impact model (or provided basic ITS model structure)	59	49.2
Use of lag period	34	28.4
Sensitivity analyses	15	12.5
Reported statistical software used	97	80.8

ACF, autocorrelation function; ARIMA, autoregressive integrated moving average; GEE, generalised estimating equation; ITS, interrupted time series; PACF, partial autocorrelation function; QI, quality improvement.;

the ITS analyses. Of concern, the risk of bias was high or very high in about three-quarters of included studies and did not appear to have changed significantly over time.

While ITS is arguably one of the strongest quasi-experimental designs available, ITS studies may be subjected to several threats to internal validity.[10 18 22] These internal validity threats can often be minimised by adding a control time series group to ITS where feasible.[10 18 22 24] However, this review found that a control group was used in less than 20% of included studies, with a location-based control being the most commonly used. When QI interventions were multisite and implemented at different times, approaches used to assign control were not often discussed clearly. Most of studies included in this review were from high-income countries; however, with expanding routine health information systems in LMICs, there is opportunity to conduct well-designed ITS studies for evaluating healthcare interventions including QI in these settings.[25 163]

Consistent with previous systematic reviews, this review highlights variations in methodological considerations (or application) and reporting of ITS.[19–21] For example,

Table 4 Reporting of ITS study results and interpretation

Results	n	%
Participants		
► Characteristics in each study group	84	70
► Flow diagram of study participant selection	11	9.2
Outcomes		
► Reported all outcomes examined over the study period)	105	87.5
► Report the average, minimum and maximum number of outcomes across time intervals	22	18.3
► Reported level/trend changes	107	89.2
► Report absolute and/or relative changes and their significance, eg, clinical, policy and statistical	35	29.2
► Report CI or SE	98	81.7
► Graphical figures to display results	111	92.5
► Fitted lines (trend)	64	57.6
► Counterfactual lines	14	12.6
► Used time lag and showed it on figure in results	23	67.6
► Results of sensitivity analyses if relevant	3	20
Interpretation		
► Key results	119	99.2
► Context (related to possible confounding)	113	97.4
► Relevant co-interventions	56	46.7
► Stability of the participant characteristics over time	24	20
► Stability of outcome coding over time	28	23.3
Limitations of the study		
► Discussion of limitations of the study	114	95
► Data variability/appropriateness of number data points	31	25.8
► Discussion direction/magnitude of any potential bias	48	42.1

Jandoc *et al*'s (2015) systematic review of application of ITS in drug utilisation research found that about three-quarters of studies examined drug utilisation over monthly intervals, and two-thirds used segmented regression techniques.[20] They also found that key methodological considerations in ITS including seasonality and non-stationarity were considered in less than half of studies.[20] However, autocorrelation was considered in two-thirds of studies included in Jandoc *et al*'s review unlike in our review where it was considered in only about half of studies.[20]

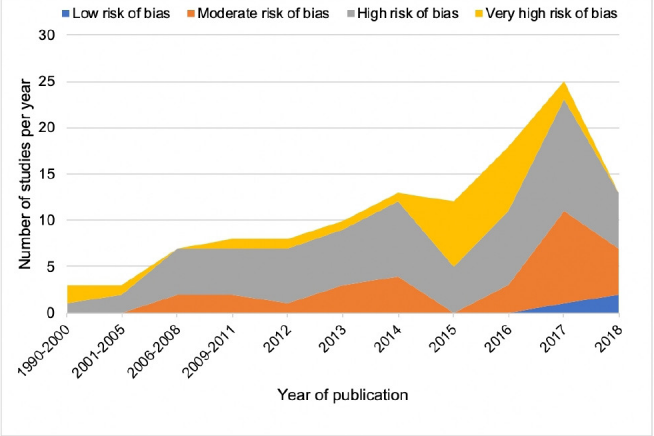

Figure 5 Number and risk of bias of included studies over time. Our literature search end date was June 2018, and as such, studies published in 2018 that were captured by our search strategy may not have been representative of all studies published in 2018.

Our findings showed that considerations of autocorrelation, seasonality and non-stationarity were more likely to be reported when ARIMA was used compared with segmented regression, which mirrors prior reviews in other subject areas.[20] Similar to this review, a more recent review by Huddson *et al* (2019) examined characteristics and reporting of single ITS in healthcare in studies published in 2015 and found that only about half of studies considered autocorrelation, 28% considered seasonality and 10% considered non-stationarity.[21] Consistent with previous reviews, this review showed that studies used various statistical approaches that may have different effect on study results, highlighting a need for further research on how various statistical analyses used in ITS can affect intervention estimates.[20 21]

While the number of ITS studies has increased over time, the conduct and reporting of ITS studies still leaves room for improvement. Consistent with previous reviews, variations in methodological considerations and reporting of ITS remain a concern, warranting a need to develop and reinforce formal reporting guidelines to improve the application of ITS in the evaluation of

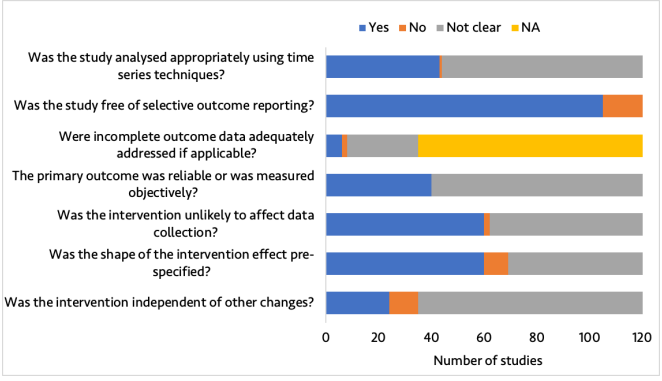

Figure 6 Summary of quality assessment of included studies. NA, not applicable. In this study, NA refers to studies with complete data.

effectiveness of healthcare interventions. Jandoc *et al* (2015) proposed methodological and reporting recommendations for ITS studies that built on STROBE, and Bernal (2018) proposed a framework for enhanced reporting of ITS framework that can be adapted to help improve design and reporting of ITS in the evaluation of QI.[20 23] Additionally, the Cochrane EPOC standard criteria for assessing risk of bias for ITS should be used to improve the design of ITS in the evaluation of QI to reduce risk of bias.[17 27] Although these EPOC standard criteria do not provide guidance on how to account for key methodological considerations such as autocorrelation, non-stationarity, seasonality and multiple interventions, there are other resources that are accessible to QI evaluators to help conduct ITS analyses appropriately, such as Wagner *et al*'s seminal paper on segmented regression analysis of ITS, Bernal *et al*'s tutorial on use of ITS in the evaluation of public health interventions and Saeed *et al*'s work on use of segmented generalised mixed effects models to evaluate health outcomes.[18 26 31]

This methodological systematic review has a number of limitations that need to be highlighted. First, our results are based on what were reported (or not) as performed in eligible studies; we did not reanalyse reported time series data to check whether the statistical analyses undertaken in these studies were performed appropriately. As such, it is possible that we underestimated the risk of bias, especially as related to whether the study data were analysed appropriately using time series techniques—one of the seven EPOC standard criteria for assessing risk of bias for ITS studies. Ramsay and colleagues reanalysed ITS studies data from two systematic reviews and found that approximately 50% of studies that had reported a significant difference were not actually significant when ITS methods were applied appropriately.[17] Second, although we did not apply any restrictions to our search strategy, we could still have missed some relevant studies. Lastly, while the recommended conventional double screening was employed in this study, data extraction was done by a single researcher.[164] However, a validation on a random subsample consisting of 10% of eligible studies found a very high agreement, making the data extraction less of a concern.

In conclusion, while use of ITS in the evaluation of health system QI interventions has increased considerably over the past decade, the quality of ITS studies still shows room for improvement. Moreover, variations in methodological considerations and reporting of ITS in QI research remain a concern, warranting a need to develop and reinforce formal reporting guidelines to improve application of ITS.

Author affiliations
[1]Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
[2]Centre for Health Services and Policy Research, School of Population and Public Health, The University of British Columbia, Vancouver, British Columbia, Canada
[3]School of Public Health, University of Rwanda, Kigali, Rwanda
[4]School of Population and Public Health, The University of British Columbia, Vancouver, British Columbia, Canada
[5]HIV/STI Surveillance Research Centre, and WHO Collaborating Centre for HIV Surveillance, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran
[6]Collaboration for Outcomes Research and Evaluation, Faculty of Pharmaceutical Sciences, The University of British Columbia, Vancouver, British Columbia, Canada
[7]Center for Health Evaluation and Outcome Sciences, Providence Health Research Institute, Vancouver, British Columbia, Canada

Twitter Celestin Hategeka @DrHategeka and Michael R Law @myclaw

Acknowledgements Dr Hategeka received support through a Vanier Canada Graduate Scholarship and a Banting Postdoctoral Fellowship from the Canadian Institutes of Health Research. Dr Karamouzian is supported by a Vanier Canada Graduate Scholarship and a Pierre Elliott Trudeau Foundation Doctoral Scholarship. Dr Law received salary support through a Canada Research Chair in Access to Medicines and a Michael Smith Foundation for Health Research Scholar Award.

Contributors CH conceived the idea and designed the study. MRL provided guidance throughout the research process. CH, MK and MRL contributed to the search strategy. CH and HR performed screening of titles, abstracts and full texts. CH and HR extracted data from eligible studies and performed risk of bias assessment. CH analysed the data and drafted the manuscript. All authors reviewed the manuscript and contributed to the revision of the manuscript and approved its final version.

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

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD
Celestin Hategeka http://orcid.org/0000-0001-7808-4652

REFERENCES

1 Kruk ME, Gage AD, Arsenault C, *et al*. High-Quality health systems in the sustainable development goals era: time for a revolution. *Lancet Glob Health* 2018;6:e1196–252.
2 National Academies of Sciences E and Medicine. *Crossing the global quality chasm: improving health care worldwide*. National Academies Press, 2018.
3 World Health Organization. *Delivering quality health services: a global imperative for universal health coverage*, 2018.
4 Shojania K, McDonald K, Wachter R, *et al*. *Closing the quality gap: a critical analysis of quality improvement strategies.(Vol. 1: Series overview and methodology). Technical Reviews, No. 9.1*. Rockville, MD: Agency for Healthcare Research and Quality, 2004.
5 Tricco AC, Ivers NM, Grimshaw JM, *et al*. Effectiveness of quality improvement strategies on the management of diabetes: a systematic review and meta-analysis. *Lancet* 2012;379:2252–61.
6 Hategeka C. *Health system performance and impact of quality improvement interventions for maternal, newborn and child health in Rwanda*. University of British Columbia, 2020.

7 Penfold RB, Zhang F. Use of interrupted time series analysis in evaluating health care quality improvements. *Acad Pediatr* 2013;13:S38–44.
8 Fretheim A, Tomic O. Statistical process control and interrupted time series: a golden opportunity for impact evaluation in quality improvement. *BMJ Qual Saf* 2015;24:748–52.
9 Lagarde M. How to do (or not to do). Assessing the impact of a policy change with routine longitudinal data. *Health Policy Plan* 2012;27:76–83.
10 Soumerai SB, Ceccarelli R, RJJogim K. *False dichotomies and health policy research designs: randomized trials are not always the answer*. , 2017: 32, 204–9.
11 Tugwell P, Knottnerus JA, McGowan J, *et al*. *Big-5 quasi-experimental designs*. , 2017: 89, 1–3.
12 Victora CG, Habicht J-P, Bryce J. Evidence-Based public health: moving beyond randomized trials. *Am J Public Health* 2004;94:400–5.
13 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215–8.
14 Ogrinc G, Mooney SE, Estrada C, *et al*. The Squire (standards for quality improvement reporting excellence) guidelines for quality improvement reporting: explanation and elaboration. *Quality and Safety in Health Care* 2008;17:i13–32.
15 Group EPOoC. *What study designs should be included in an EPOC review and what should they be called*, 2014.
16 Ewusie JE. *Improved methods for interrupted time series analysis useful when outcomes are aggregated: accounting for heterogeneity across patients and healthcare settings*, 2019.
17 Ramsay CR, Matowe L, Grilli R, *et al*. Interrupted time series designs in health technology assessment: lessons from two systematic reviews of behavior change strategies. *Int J Technol Assess Health Care* 2003;19:613–23.
18 Wagner AK, Soumerai SB, Zhang F, *et al*. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther* 2002;27:299–309.
19 Polus S, Pieper D, Burns J, *et al*. Heterogeneity in application, design, and analysis characteristics was found for controlled before-after and interrupted time series studies included in Cochrane reviews. *J Clin Epidemiol* 2017;91:56–69.
20 Jandoc R, Burden AM, Mamdani M, *et al*. Interrupted time series analysis in drug utilization research is increasing: systematic review and recommendations. *J Clin Epidemiol* 2015;68:950–6.
21 Hudson J, Fielding S, Ramsay CR. Methodology and reporting characteristics of studies using interrupted time series design in healthcare. *BMC Med Res Methodol* 2019;19:137.
22 Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*, 2002.
23 López Bernal J. *The use of interrupted time series for the evaluation of public health interventions*. London School of Hygiene & Tropical Medicine, 2018.
24 Lopez Bernal J, Cummins S, Gasparrini A. The use of controls in interrupted time series studies of public health interventions. *Int J Epidemiol* 2018;47:2082–93.
25 Wagenaar BH, Sherr K, Fernandes Q, *et al*. Using routine health information systems for well-designed health evaluations in low- and middle-income countries. *Health Policy Plan* 2016;31:129–35.
26 Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol* 2017;46:348–55.
27 EPOC. *Epoc methods paper: including interrupted time series (its) designs in a EPOC review*, 1998.
28 Turner SL, Karahalios A, Forbes AB, *et al*. Creating effective interrupted time series graphs: review and recommendations. *Res Synth Methods* 2020:jrsm.1435.
29 Bernal JL, Cummins S, AJljoe G. *Letter to the editor difference in difference, controlled interrupted time series and synthetic controls*. 1, 2019.
30 Baicker K, Svoronos T. *Testing the validity of the single interrupted time series design. Report No. 0898-2937*. National Bureau of Economic Research, 2019.
31 Saeed S, Moodie EEM, Strumpf EC, *et al*. Segmented generalized mixed effect models to evaluate health outcomes. *Int J Public Health* 2018;63:547–51.
32 Zhang F, Wagner AK, Ross-Degnan D. Simulation-Based power calculation for designing interrupted time series analyses of health policy interventions. *J Clin Epidemiol* 2011;64:1252–61.
33 Honaker J, King G. What to do about missing values in time-series cross-section data. *Am J Pol Sci* 2010;54:561–81.
34 Moritz S, Bartz-Beielstein T. imputeTS: time series missing value imputation in R. *R J* 2017;9:207.
35 pp.Allison PD. *Handling missing data by maximum likelihood*. SAS global forum, 2012: 1038–1021.
36 Durbin J, Watson GS. *Testing for serial correlation in least squares regression.I. breakthroughs in statistics*. Springer, 1992: 237–59.
37 Bhaskaran K, Gasparrini A, Hajat S, *et al*. Time series regression studies in environmental epidemiology. *Int J Epidemiol* 2013;42:1187–95.
38 Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc* 1979;74:427–31.
39 Hategeka C, Karamouzian M, Law MR. *Use of interrupted time series analysis in the evaluation of health care quality improvement interventions: a methodological systematic review*. PROSPERO, 2018.
40 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9.
41 Cochrane Effective Practice and Organisation of Care (EPOC). *EPOC resources for review authors*, 2017.
42 Rowe AK, Rowe SY, Peters DH, *et al*. Effectiveness of strategies to improve health-care provider practices in low-income and middle-income countries: a systematic review. *Lancet Glob Health* 2018;6:e1163–75.
43 Shenoy AG, Begley CE, Revere L, *et al*. Innovating patient care delivery: DSRIP's interrupted time series analysis paradigm. *Healthcare*2018.
44 Sheibani R, Sheibani M, Heidari-Bakavoli A, *et al*. The effect of a clinical decision support system on improving adherence to guideline in the treatment of atrial fibrillation: an interrupted time series study. *J Med Syst* 2018;42.
45 Savage T, Hodge DE, Pickard K, *et al*. Sustained reduction and prevention of neonatal and pediatric central line-associated bloodstream infection following a Nurse-Driven quality improvement initiative in a pediatric facility. *Journal of the Association for Vascular Access* 2018;23:30–41.
46 Patel S, Carmichael JM, Taylor JM, *et al*. Evaluating the impact of a clinical decision support tool to reduce chronic opioid dose and decrease risk classification in a veteran population. *Ann Pharmacother* 2018;52:325–31.
47 Mills AM, Ip IK, Langlotz CP, *et al*. Clinical decision support increases diagnostic yield of computed tomography for suspected pulmonary embolism. *Am J Emerg Med* 2018;36:540–4.
48 Mehtsun WT, Papanicolas I, Zheng J, *et al*. National trends in readmission following inpatient surgery in the hospital readmissions reduction program era. *Ann Surg* 2018;267:599–605.
49 Karami Matin B, Hajizadeh M, Najafi F, *et al*. The impact of health sector evolution plan on hospitalization and cesarean section rates in Iran: an interrupted time series analysis. *Int J Qual Health Care* 2018;30:75–9.
50 Liu B, Moore JE, Almaawiy U, *et al*. Outcomes of mobilisation of vulnerable elders in Ontario (move on): a multisite interrupted time series evaluation of an implementation intervention to increase patient mobilisation. *Age Ageing* 2018;47:112–9.
51 Kim J-Y, Kim SJ, Nam CM, *et al*. Changes in prescription pattern, pharmaceutical expenditure and quality of care after introduction of reimbursement restriction in diabetes in Korea. *Eur J Public Health* 2018;28:209–14.
52 Herbst T, Foerster J, Emmert M. The impact of pay-for-performance on the quality of care in ophthalmology: empirical evidence from Germany. *Health Policy* 2018;122:667–73.
53 Chua CC, Hutchinson A, Tacey M, *et al*. A physician targeted intervention improves prescribing in chronic heart failure in general medical units. *BMC Health Serv Res* 2018;18:206.
54 Chen I-L, Lee C-H, Su L-H, *et al*. Effects of implementation of an online comprehensive antimicrobial-stewardship program in ICUs: a longitudinal study. *J Microbiol Immunol Infect* 2018;51:55–63.
55 Barker AK, Krasity B, Musuuza J, *et al*. Screening for asymptomatic Clostridium difficile among bone marrow transplant patients: a mixed-methods study of intervention effectiveness and feasibility. *Infect Control Hosp Epidemiol* 2018;39:177–85.
56 Willis ZI, Gillon J, Xu M, *et al*. Reducing antimicrobial use in an academic pediatric institution: evaluation of the effectiveness of a prospective audit with real-time feedback. *J Pediatric Infect Dis Soc* 2017;6:339–45.
57 Wattier RL, Levy ER, Sabnis AJ, *et al*. Reducing second gram-negative antibiotic therapy on pediatric oncology and hematopoietic stem cell transplantation services. *Infect Control Hosp Epidemiol* 2017;38:1039–47.
58 Wagner AD, Mugo C, Bluemer-Miroite S, *et al*. Continuous quality improvement intervention for adolescent and young adult HIV

testing services in Kenya improves HIV knowledge. *AIDS* 2017;31 Suppl 3:S243–52.
59 Theobald CN, Choma NN, Ehrenfeld JM, *et al*. Effect of a handover tool on efficiency of care and mortality for interhospital transfers. *J Hosp Med* 2017;12:23–8.
60 Taylor JE, McDonald SJ, Earnest A, *et al*. A quality improvement initiative to reduce central line infection in neonates using checklists. *Eur J Pediatr* 2017;176:639–46.
61 Sun J, Lin Q, Zhao P, *et al*. Reducing waiting time and raising outpatient satisfaction in a Chinese public tertiary General hospital-an interrupted time series study. *BMC Public Health* 2017;17.
62 Sun BC, Lupulescu-Mann N, Charlesworth CJ, *et al*. Impact of Hospital "Best Practice" Mandates on Prescription Opioid Dispensing After an Emergency Department Visit. *Acad Emerg Med* 2017;24:905–13.
63 Song I, Shin HN, Shin J-Y. Decrease in use of contraindicated drugs with automated alerts in children. *Pediatr Int* 2017;59:720–6.
64 Skinner S, Davies-Tuck M, Wallace E, *et al*. Perinatal and maternal outcomes after training residents in forceps before vacuum instrumental birth. *Obstet Gynecol* 2017;130:151–8.
65 Scott JW, Nyinawankusi Jeanne D'Arc, Enumah S, *et al*. Improving prehospital trauma care in Rwanda through continuous quality improvement: an interrupted time series analysis. *Injury* 2017;48:1376–81.
66 Pratt NL, Kalisch Ellett LM, Sluggett JK, *et al*. Use of proton pump inhibitors among older Australians: national quality improvement programmes have led to sustained practice change. *Int J Qual Health Care* 2017;29:75–82.
67 Ouldali N, Bellêttre X, Milcent K, *et al*. Impact of implementing national guidelines on antibiotic prescriptions for acute respiratory tract infections in pediatric emergency departments: an interrupted time series analysis. *Clin Infect Dis* 2017;65:1469–76.
68 Muzigaba M, Kigozi G, Puoane T. Short-Term and sustained effects of a health system strengthening intervention to improve mortality trends for paediatric severe malnutrition in rural South African hospitals: an interrupted time series design. *S Afr J CH* 2017;11:38–45.
69 Miwa S, Visintainer P, Engelman R, *et al*. Effects of an ambulation orderly program among cardiac surgery patients. *Am J Med* 2017;130:1306–12.
70 Lin C-W, Wang C-H, Huang W-I, *et al*. Impact of Safety-Related regulations on codeine use in children: a quasi-experimental study using Taiwan's National health insurance research database. *Drug Saf* 2017;40:615–27.
71 Li L, Fortin E, Tremblay C, *et al*. Hospital-Acquired methicillin-resistant Staphylococcus aureus bloodstream infections in Québec: impact of guidelines. *Infect Control Hosp Epidemiol* 2017;38:840–7.
72 Lapointe-Shaw L, Mamdani M, Luo J, *et al*. Effectiveness of a financial incentive to physicians for timely follow-up after hospital discharge: a population-based time series analysis. *CMAJ* 2017;189:E1224–9.
73 Horton DJ, Yarbrough PM, Wanner N, *et al*. Improving physician communication with patients as measured by HCAHPS using a standardized communication model. *Am J Med Qual* 2017;32:617–24.
74 Erickson KF, Winkelmayer WC, Chertow GM, *et al*. Hemodialysis hospitalizations and readmissions: the effects of payment reform. *Am J Kidney Dis* 2017;69:237–46.
75 Sarin E, Kole SK, Patel R, Enisha S, Rachana P, *et al*. Evaluation of a quality improvement intervention for obstetric and neonatal care in selected public health facilities across six states of India. *BMC Pregnancy Childbirth* 2017;17:134.
76 Ehrenfeld JM, Wanderer JP, Terekhov M, *et al*. A perioperative systems design to improve intraoperative glucose monitoring is associated with a reduction in surgical site infections in a diabetic patient population. *Anesthesiology* 2017;126:431–40.
77 Charani E, Gharbi M, Moore LSP, *et al*. Effect of adding a mobile health intervention to a multimodal antimicrobial stewardship programme across three teaching hospitals: an interrupted time series study. *J Antimicrob Chemother* 2017;72:1825–31.
78 Berrevoets MAH, Pot JHLW, Houterman AE, *et al*. An electronic trigger tool to optimise intravenous to oral antibiotic switch: a controlled, interrupted time series study. *Antimicrob Resist Infect Control* 2017;6:81.
79 Bendzsak AM, Baxter NN, Darling GE, *et al*. Regionalization and outcomes of lung cancer surgery in Ontario, Canada. *J Clin Oncol* 2017;35:2772–80.
80 Angelidou A, Bell K, Gupta M, *et al*. Implementation of a guideline to decrease use of Acid-Suppressing medications in the NICU. *Pediatrics* 2017;140:e20171715–10.
81 Vercheval C, Gillet M, Maes N, *et al*. Quality of documentation on antibiotic therapy in medical records: evaluation of combined interventions in a teaching hospital by repeated point prevalence survey. *Eur J Clin Microbiol Infect Dis* 2016;35:1495–500.
82 Utnes Høgli J, Hennie Garcia B, Skjold F, *et al*. An audit and feedback intervention study increased adherence to antibiotic prescribing guidelines at a Norwegian Hospital. *BMC Infectious Diseases* 2016;16:1–11.
83 Tzeng I-S, Liu S-H, Chen K-F, *et al*. Impact of performance grading on annual numbers of acute myocardial infarction-associated emergency department visits in Taiwan: results of segmented regression analysis. *Medicine* 2016;95:e4937.
84 Thompson W, Hogel M, Li Y, *et al*. Effect of a proton pump inhibitor deprescribing guideline on drug usage and costs in long-term care. *J Am Med Dir Assoc* 2016;17
85 Szatkowski L, Aveyard P. Provision of smoking cessation support in UK primary care: impact of the 2012 QOF revision. *Br J Gen Pract* 2016;66:e10–15.
86 Stewart TC, Clark A, Gilliland J, *et al*. Home safe home: evaluation of a childhood home safety program. *J Trauma Acute Care Surg* 2016;81:533–40.
87 Prabhakaran S, Lee J, O'Neill K, *et al*. Regional learning collaboratives produce rapid and sustainable improvements in stroke thrombolysis times. *Circ Cardiovasc Qual Outcomes* 2016;9:585–92.
88 Sayah A, Lai-Becker M, Kingsley-Rocker L, *et al*. Emergency department expansion versus patient flow improvement: impact on patient experience of care. *J Emerg Med* 2016;50:339–48.
89 Rutman L, Atkins RC, Migita R, *et al*. Modification of an established pediatric asthma pathway improves evidence-based, efficient care. *Pediatrics* 2016;138:peds.2016-1248.
90 Rhodes D, Cheng AC, McLellan S, *et al*. Reducing Staphylococcus aureus bloodstream infections associated with peripheral intravenous cannulae: successful implementation of a care bundle at a large Australian health service. *J Hosp Infect* 2016;94:86–91.
91 Nguyen-Ha P-T, Howrie D, Crowley K, *et al*. A quality assessment of a collaborative model of a pediatric antimicrobial stewardship program. *Pediatrics* 2016;137:e1–8.
92 Parikh K, Hall M, Blaschke AJ, *et al*. Aggregate and hospital-level impact of national guidelines on diagnostic resource utilization for children with pneumonia at children's hospitals. *J Hosp Med* 2016;11:317–23.
93 Lee KR, Bagga B, Arnold SR. Reduction of broad-spectrum antimicrobial use in a tertiary children's Hospital post antimicrobial stewardship program guideline implementation. *Pediatr Crit Care Med* 2016;17:187–93.
94 Kim J-S, Chung YK, Lee SS, *et al*. Effect of daily chlorhexidine bathing on the acquisition of methicillin-resistant Staphylococcus aureus in a medical intensive care unit with methicillin-resistant S aureus endemicity. *Am J Infect Control* 2016;44:1520–5.
95 Singh K, Brodish P, Speizer I, *et al*. Can a quality improvement project impact maternal and child health outcomes at scale in northern Ghana? *Health Res Policy Syst* 2016;14:45.
96 Huh K, Chung DR, Park HJ, *et al*. Impact of monitoring surgical prophylactic antibiotics and a computerized decision support system on antimicrobial use and antimicrobial resistance. *Am J Infect Control* 2016;44:e145–52.
97 Harris AHS, Bowe T, Hagedorn H, *et al*. Multifaceted academic detailing program to increase pharmacotherapy for alcohol use disorder: interrupted time series evaluation of effectiveness. *Addict Sci Clin Pract* 2016;11:15.
98 Boel J, Andreasen V, Jarløv JO, *et al*. Impact of antibiotic restriction on resistance levels of Escherichia coli: a controlled interrupted time series study of a hospital-wide antibiotic stewardship programme. *J Antimicrob Chemother* 2016;71:2047–51.
99 Williams DJ, Edwards KM, Self WH, *et al*. Antibiotic choice for children hospitalized with pneumonia and adherence to national guidelines. *Pediatrics* 2015;136:44–52.
100 Milder EA, Rizzi MD, Morales KH, *et al*. Impact of a new practice guideline on antibiotic use with pediatric tonsillectomy. *JAMA Otolaryngol Head Neck Surg* 2015;141:410–6.
101 Lion KC, Ebel BE, Rafton S, *et al*. Evaluation of a quality improvement intervention to increase use of telephonic interpretation. *Pediatrics* 2015;135:e709–16.
102 Laverty AA, Elkin SL, Watt HC, *et al*. Impact of a COPD discharge care bundle on readmissions following admission with acute exacerbation: interrupted time series analysis. *PLoS One* 2015;10:e0116187.
103 Ir P, Korachais C, Chheng K, *et al*. Boosting facility deliveries with results-based financing: a mixed-methods evaluation of

the government midwifery incentive scheme in Cambodia. *BMC Pregnancy Childbirth* 2015;15:170.
104 Gonzalo JD, Kuperman EF, Chuang CH, *et al*. Impact of an overnight internal medicine academic hospitalist program on patient outcomes. *J Gen Intern Med* 2015;30:1795–802.
105 Gallagher N, Cardwell C, Hughes C, *et al*. Increase in the pharmacological management of type 2 diabetes with pay-for-performance in primary care in the UK. *Diabet Med* 2015;32:62–8.
106 Dunne RM, Ip IK, Abbett S, *et al*. Effect of evidence-based clinical decision support on the use and yield of CT pulmonary angiographic imaging in hospitalized patients. *Radiology* 2015;276:167–74.
107 Dik J-WH, Hendrix R, Lo-Ten-Foe JR, *et al*. Automatic day-2 intervention by a multidisciplinary antimicrobial stewardship-team leads to multiple positive effects. *Front Microbiol* 2015;6:546.
108 Cha WC, Song KJ, Cho JS, *et al*. The long-term effect of an independent capacity protocol on emergency department length of stay: a before and after study. *Yonsei Med J* 2015;56:1428–36.
109 Borde JP, Litterst S, Ruhnke M, *et al*. Implementing an intensified antibiotic stewardship programme targeting cephalosporin and fluoroquinolone use in a 200-bed community hospital in Germany. *Infection* 2015;43:45–50.
110 Balicer RD, Hoshen M, Cohen-Stavi C, *et al*. Sustained reduction in health disparities achieved through targeted quality improvement: one-year follow-up on a three-year intervention. *Health Serv Res* 2015;50:1891–909.
111 Yang P, Cao Y, Liu D, *et al*. The effect of electronic medical record application on the length of stay in a Chinese General Hospital: a department- and disease-focused interrupted time-series study. *J Med Syst* 2014;38.
112 Wang H-Y, Chiu C-H, Huang C-T, *et al*. Blood culture-guided de-escalation of empirical antimicrobial regimen for critical patients in an online antimicrobial stewardship programme. *Int J Antimicrob Agents* 2014;44:520–7.
113 Self WH, Mickanin J, Grijalva CG, *et al*. Reducing blood culture contamination in community hospital emergency departments: a multicenter evaluation of a quality improvement intervention. *Acad Emerg Med* 2014;21:274–82.
114 Sales AE, Schalm C, Baylon MAB, *et al*. Data for improvement and clinical excellence: report of an interrupted time series trial of feedback in long-term care. *Implement Sci* 2014;9:161.
115 Ranzani OT, Simpson ES, Augusto TB, *et al*. Evaluation of a minimal sedation protocol using ICU sedative consumption as a monitoring tool: a quality improvement multicenter project. *Crit Care* 2014;18:580.
116 Parikh K, Hall M, Teach SJ. Bronchiolitis management before and after the Aap guidelines. *Pediatrics* 2014;133:e1–7.
117 Devkaran S, O'Farrell PN. The impact of hospital accreditation on clinical documentation compliance: a life cycle explanation using interrupted time series analysis. *BMJ Open* 2014;4:e005240.
118 Derde LPG, Cooper BS, Goossens H, *et al*. Interventions to reduce colonisation and transmission of antimicrobial-resistant bacteria in intensive care units: an interrupted time series study and cluster randomised trial. *Lancet Infect Dis* 2014;14:31–9.
119 Battersby C, Santhakumaran S, Upton M, *et al*. The impact of a regional care bundle on maternal breast milk use in preterm infants: outcomes of the East of England quality improvement programme. *Arch Dis Child Fetal Neonatal Ed* 2014;99:F395–401.
120 Andersson Hagiwara M, Suserud B-O, Andersson-Gäre B, *et al*. The effect of a computerised decision support system (CDSS) on compliance with the prehospital assessment process: results of an interrupted time-series study. *BMC Med Inform Decis Mak* 2014;14:70.
121 Amin AN, Hofmann H, Owen MM, *et al*. *Reduce readmissions with Service-Based care management. professional case management*. Baltimore, Maryland: Lippincott Williams & Wilkins, 2014: 255–64.
122 Aldeyab MA, Scott MG, Kearney MP, *et al*. Impact of an enhanced antibiotic stewardship on reducing methicillin-resistant Staphylococcus aureus in primary and secondary healthcare settings. *Epidemiol Infect* 2014;142:494–500.
123 Akenroye AT, Baskin MN, Samnaliev M, *et al*. Impact of a bronchiolitis guideline on ED resource use and cost: a segmented time-series analysis. *Pediatrics* 2014;133:e227–34.
124 Sarma JB, Marshall B, Cleeve V, *et al*. Impact of universal screening on MRSA bacteremias in a single acute NHS organisation (2006-12): interrupted time-series analysis. *Antimicrob Resist Infect Control* 2013;2:2.
125 Roughead EE, Kalisch Ellett LM, Ramsay EN, *et al*. Bridging evidence-practice gaps: improving use of medicines in elderly Australian veterans. *BMC Health Serv Res* 2013;13:514.
126 Patel PR, Yi SH, Booth S, *et al*. Bloodstream infection rates in outpatient hemodialysis facilities participating in a collaborative prevention effort: a quality improvement report. *Am J Kidney Dis* 2013;62:322–30.
127 McAlister FA, Bakal JA, Kaul P, *et al*. Changes in heart failure outcomes after a province-wide change in health service provision a natural experiment in Alberta, Canada. *Circ Heart Fail* 2013;6:76-82.
128 Elsaid K, Truong T, Monckeberg M, *et al*. Impact of electronic chemotherapy order forms on prescribing errors at an urban medical center: results from an interrupted time-series analysis. *Int J Qual Health Care* 2013;25:656–63.
129 Desai SP, Lu B, Szent-Gyorgyi LE, *et al*. Increasing pneumococcal vaccination for immunosuppressed patients: a cluster quality improvement trial. *Arthritis Rheum* 2013;65:39–47.
130 Cairns KA, Jenney AWJ, Abbott IJ, *et al*. Prescribing trends before and after implementation of an antimicrobial stewardship program. *Med J Aust* 2013;198:262–6.
131 Butcher BW, Vittinghoff E, Maselli J, *et al*. Impact of proactive rounding by a rapid response team on patient outcomes at an academic medical center. *J Hosp Med* 2013;8:7–12.
132 Andersen SE, Knudsen JD, Bispebjerg Intervention Group. A managed multidisciplinary programme on multi-resistant Klebsiella pneumoniae in a Danish university hospital. *BMJ Qual Saf* 2013;22:907–15.
133 Aiken AM, Wanyoro AK, Mwangi J, *et al*. Changing use of surgical antibiotic prophylaxis in Thika Hospital, Kenya: a quality improvement intervention with an interrupted time series design. *PLoS One* 2013;8:e78942.
134 Yeo C-L, Chan DS-G, Earnest A, *et al*. Prospective audit and feedback on antibiotic prescription in an adult hematology-oncology unit in Singapore. *Eur J Clin Microbiol Infect Dis* 2012;31:583–90.
135 Webster PD, Sibanyoni M, Malekutu D, *et al*. Using quality improvement to accelerate highly active antiretroviral treatment coverage in South Africa. *BMJ Qual Saf* 2012;21:315–24.
136 Slekovec C, Leroy J, Vernaz-Hegi N, *et al*. Impact of a region wide antimicrobial stewardship guideline on urinary tract infection prescription patterns. *Int J Clin Pharm* 2012;34:325–9.
137 Morgan DJ, Pineles L, Shardell M, *et al*. Automated hand hygiene count devices may better measure compliance than human observation. *Am J Infect Control* 2012;40:955–9.
138 Magedanz L, Silliprandi EM, dos Santos RP, *et al*. Impact of the pharmacist on a multidisciplinary team in an antimicrobial stewardship program: a quasi-experimental study. *Int J Clin Pharm* 2012;34:290–4.
139 Kinsman L, Buykx P, Cant R, *et al*. The FIRST ^{2} ACT simulation program improves nursing practice in a rural Australian hospital. *Australian Journal of Rural Health* 2012;20:270–4.
140 Elligsen M, Walker SAN, Pinto R, *et al*. Audit and feedback to reduce broad-spectrum antibiotic use among intensive care unit patients: a controlled interrupted time series analysis. *Infect Control Hosp Epidemiol* 2012;33:354–61.
141 Alshamsan R, Lee JT, Majeed A, *et al*. Effect of a UK pay-for-performance program on ethnic disparities in diabetes outcomes: interrupted time series analysis. *Ann Fam Med* 2012;10:228–34.
142 Serumaga B, Ross-Degnan D, Avery AJ, *et al*. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *BMJ* 2011;342:d108.
143 Haynes K, Linkin DR, Fishman NO, *et al*. Effectiveness of an information technology intervention to improve prophylactic antibacterial use in the postoperative period. *J Am Med Inform Assoc* 2011;18:164–8.
144 Gadzhanova S, Roughead EE, Loukas K, *et al*. Improvement in metformin and insulin utilisation in the Australian veteran population associated with quality use of medicines intervention programs. *Pharmacoepidemiol Drug Saf* 2011;20:359–65.
145 Borer A, Eskira S, Nativ R, *et al*. A multifaceted intervention strategy for eradication of a hospital-wide outbreak caused by carbapenem-resistant Klebsiella pneumoniae in southern Israel. *Infect Control Hosp Epidemiol* 2011;32:1158–65.
146 Helder OK, Brug J, Looman CWN, *et al*. The impact of an education program on hand hygiene compliance and nosocomial infection incidence in an urban neonatal intensive care unit: an intervention study with before and after comparison. *Int J Nurs Stud* 2010;47:1245–52.
147 van Doormaal JE, van den Bemt PMLA, Zaal RJ, *et al*. The influence that electronic prescribing has on medication errors and preventable adverse drug events: an interrupted time-series study. *J Am Med Inform Assoc* 2009;16:816–25.

148 Gulihar A, Nixon M, Jenkins D, *et al*. Clostridium difficile in hip fracture patients: prevention, treatment and associated mortality. *Injury* 2009;40:746–51.
149 Colombet I, Sabatier B, Gillaizeau F, *et al*. Long-Term effects of a multifaceted intervention to encourage the choice of the oral route for proton pump inhibitors: an interrupted time-series analysis. *Qual Saf Health Care* 2009;18:232–5.
150 Audit NPT. Impact of NICE guidance on rates of haemorrhage after tonsillectomy: an evaluation of guidance issued during an ongoing national tonsillectomy audit. *Qual Saf Health Care* 2008;17:264–8.
151 Buising KL, Thursky KA, Black JF, *et al*. Improving antibiotic prescribing for adults with community acquired pneumonia: Does a computerised decision support system achieve more than academic detailing alone?--A time series analysis. *BMC Med Inform Decis Mak* 2008;8:35.
152 Mamdani M, McNeely D, Evans G, *et al*. Impact of a fluoroquinolone restriction policy in an elderly population. *Am J Med* 2007;120:893–900.
153 Smith DH, Perrin N, Feldstein A, *et al*. The impact of prescribing safety alerts for elderly persons in an electronic medical record: an interrupted time series evaluation. *Arch Intern Med* 2006;166:1098–104.
154 Simon SR, Smith DH, Feldstein AC, *et al*. Computerized prescribing alerts and group academic detailing to reduce the use of potentially inappropriate medications in older people. *J Am Geriatr Soc* 2006;54:963–8.
155 Madaras-Kelly KJ, Remington RE, Lewis PG, *et al*. Evaluation of an intervention designed to decrease the rate of nosocomial methicillin-resistant Staphylococcus aureus infection by encouraging decreased fluoroquinolone use. *Infect Control Hosp Epidemiol* 2006;27:155–69.
156 Grégoire J-P, Moisan J, Potvin L, *et al*. Effect of drug utilization reviews on the quality of in-hospital prescribing: a quasi-experimental study. *BMC Health Serv Res* 2006;6:33.
157 van Kasteren MEE, Mannien J, Kullberg B-J, *et al*. Quality improvement of surgical prophylaxis in Dutch hospitals: evaluation of a multi-site intervention by time series analysis. *J Antimicrob Chemother* 2005;56:1094–102.
158 Poulose BK, Ray WA, Arbogast PG, *et al*. Resident work hour limits and patient safety. *Ann Surg* 2005;241:847–60.
159 Neilson EG, Johnson KB, Rosenbloom ST, *et al*. The impact of peer management on test-ordering behavior. *Ann Intern Med* 2004;141:196–204.
160 Durieux P, Nizard R, Ravaud P, *et al*. A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior. *JAMA* 2000;283:2816–21.
161 Coburn AF, Fortinsky R, McGuire C, *et al*. Effect of prospective reimbursement on nursing home costs. *Health Serv Res* 1993;28:45–68.
162 Everitt DE, Soumerai SB, Avorn J, *et al*. Changing surgical antimicrobial prophylaxis practices through education targeted at senior department leaders. *Infect Control Hosp Epidemiol* 1990;11:578–83.
163 Ashton RA, Bennett A, Yukich J, *et al*. Methodological considerations for use of routine health information system data to evaluate malaria program impact in an era of declining malaria transmission. *Am J Trop Med Hyg* 2017;97:46–57.
164 Waffenschmidt S, Knelangen M, Sieben W, *et al*. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. *BMC Med Res Methodol* 2019;19:132.