HAUTE AUTORITÉ DE SANTÉ

MEASURE
& IMPROVE THE QUALITY

METHODOLOGICAL GUIDE

Outcome indicators measured from medico-administrative databases

Method regarding development, validation and uses

This document is a translation of the original French document

June 2019

Description of the publication

Titre	Outcome indicators measured from medico-administrative databases Method regarding development, validation and uses
Method of production	Methodological report
Objective(s)	This document presents the HAS methodology for the development, validation and the conditions required for uses of outcome indicators measured from medico-administrative databases. Advantages and limitations of quality indicators based on medico-administrative data are also reported.
Targets	Users, professionals, institutions
Requester	Self- referral
Sponsor	Haute Autorité de santé (HAS)
Project management	Linda Banaei-Bouchareb, Project Manager, EvOQSS Department (Head of Department: Laetitia May-Michelangeli; Deputy Head of Department: Sandrine Morin).
Literature search	Virginie Henry, Archivist; Renée Cardoso, Assistant Archivist.
Authors	Linda Banaei-Bouchareb, Project Manager; Frédéric Capuano and Agnès Solomiac, Statisticians; Sylvaine Egea and Anne Grigorian, Project Assistants.
Declaration of interests	Members of the working group shared their public statements of interest with HAS. They can be found at https://dpi.sante.gouv.fr. They were analyzed according to the analysis grid in the HAS Declaration of Interest and Conflict of Interest Management Guide. The interests declared by the members of the working group were considered compatible with their participation in this work.
Validation	June 2019

Ce document ainsi que sa référence bibliographique sont téléchargeables sur www.has-sante.fr

Haute Autorité de santé – Service communication et information
5 avenue du Stade de France – 93218 SAINT-DENIS LA PLAINE CEDEX. Tél. : +33 (0)1 55 93 70 00

Table of contents

1. Definitions, advantages and limitations

HAS develops care quality and safety indicators (IQSSs) with the healthcare professionals involved in the provision of care, as well as patients and healthcare system users, using a validated method. The purpose of these indicators is to help improve the quality of care and safety of patients in healthcare organisations.

1.1. Definitions

A care quality and safety indicator is a tool to measure a health state, a practice or the occurrence of an event. It makes it possible to estimate the quality of care and its variations over time in a valid and reliable way. The link between the indicator and the quality of care must first be demonstrated through a literature review and/or an opinion issued by the work group (gathering expertise in clinical/care, medical information, quality management procedure, as well as patients and healthcare system users point of view).

These indicators may be classified into types – structure, process or outcome (according to the Donabedian Classification) – and can be measured using various data sources (ANAES, 2002; Shaw and Kalo, 2002; Donabedian, 1988).

An outcome indicator directly measures, after a care process, the benefits or risks generated for the patient in terms of efficacy, satisfaction, safety and cost-effectiveness. They generally require an adjustment with the variables that influence the outcome independently from the quality of care (e.g. characteristics of the population).

Outcome indicators provide a response to:

- the expectations of healthcare system users;
- the demand of healthcare professionals;
- a shared goal to improve the outcome for the patient.

They rely on the availability of professional references which define recommended practices in view of current medical knowledge, and the use of medico-administrative databases, given the lack of available nation-wide clinical information systems.

1.2. Advantages and limitations of outcome indicators

Advantages

- These outcome indicators make it possible to measure the impact of care practices and improvement actions/strategies.
- They make it possible to identify the structures and/or patient populations needing to be investigated.
- They are a key element to raise the awareness of healthcare professionals on the possible flaws in the care organisation and/or clinical practice.
- Their integrated use in quality/risk management procedures – in particular via the certification of healthcare organisations – is a good improvement driver.

Limitations

The variables that influence the outcome independently from the quality of care must be identified. They can be used for adjustment purposes if they are not linked to the quality of care, or to explain the variability of the results of the adjusted measure.

Their use for quality improvement requires:

- a patient record review;
- to analyse the care practices related to the measured outcome. Depending on the degree of correlation between outcome and practices, the outcome indicator may be more or less sensitive to changes in these practices.

1.3. Advantages and limitations of their measurement from medico-administrative databases

Advantages

The main advantages of the measurement of outcome indicators using medico-administrative databases are their automatic calculation from existing data without an additional workload for healthcare professionals, and the possible follow up over time.

Limitations

Medico-administrative databases provide access to retrospective data, such as the national consolidated Program of Medicalization of Information Systems (PMSI) data available on the Agence Technique de l'information sur l'Hospitalisation (the French technical agency for information on hospital care – ATIH) server in May of year N+1. The data are subsequently available on the SNDS (French national administrative healthcare database), linked with the data of the national health insurance inter-scheme IT system (SNIIRAM) in August of year N+1.

The validity of the indicator depends on the availability and the quality of the events coding (Januel, DREES, 2011), along with the comorbidities/risk factors used in the adjustment. Indeed, the medico-administrative databases are primarily designed for the pricing of activities, rather than to describe patient care in an exhaustive way.

"Adverse effects" have been reported regarding the use of outcome indicators measured from medico-administrative databases for public disclosure or regulation purposes (Heath et al. 2007; Gubb, 2009; HAS, 2017; Steven et al. 2012), as the healthcare organisations' efforts were then focused on optimising results for budget purposes, unrelated to the improvement of the quality of care and patient safety. Examples drawn from international experiences include the under-reporting of events (Farmer et al. 2013) and the change in admission/discharge/transfer policies, in order to select the patients with the lowest risk or to avoid the attribution of an adverse event (HAS, 2017).

The main advantages of the outcome indicators measured from medico-administrative databases are: their automatic calculation from existing data without an additional workload for healthcare professionals, the possible follow up over time, and their ability to identify the structures and/or patient populations needing to be investigated.

Their current drawbacks are their limitation due to the fact that their metrological qualities depend on coding practices, and the relatively long amount of time required for the data to be made available. The limitations imply the need for vigilance in the use of this type of indicator for purposes other than the internal management of the quality of care.

2. Context

In France, the DGOS (Directorate General of Health Care Provision) and HAS collaborated for the set-up of indicators derived from research projects to develop care quality and safety indicators (IQSSs).

The work done since 2015 for the development of outcome indicators measured from the PMSI is in line with the scientific adaptation of patient safety indicators (PSIs) to the French context (Januel, DREES, 2011, Januel, 2011). This work includes the drafting of PSI development, use and dissemination guidelines by Consortium Loire-atlantique Aquitaine Rhône-Alpes pour la production d'indicateurs en sanTÉ (CLARTE) (Le Pogam et al. CLARTE, 2012), work on the validation of PSIs (Rapport CLARTE 2010-2013), and scientific studies (2007-2017) on the development of mortality measurement in France (HAS. Indicateurs de mortalité hospitalière, 2017; Lamarche-Vadel et al. AMPHI, 2014; Januel, DREES, 2011). The resulting HAS method validates outcome indicators based on published international criteria (Januel, DREES, 2011; Le Pogam CLARTE, 2012; Romano, 2009; Davies et al. AHRQ, 2001; Taffé et al. 2012).

This document describes the strategy put in place by HAS for the development, validation and use of outcome indicators measured from medico-administrative databases. This strategy is based on a method designed to improve the quality and safety of care and the outcome for patients (see Figure 1 and Table 2).

The method described in this report was used for the development and testing of outcome indicators measured from the PMSI database from 2015 to 2018. For the purpose of internal management of the quality and safety of care, these indicators measure the following:

- thromboembolic events after total hip or knee replacement surgery (ETE-ORTHO);
- surgical site infections 3 months after total hip or knee replacement (ISO-ORTHO);
- rehospitalisations 1 to 3 days after outpatient surgery.

This method has been adapted for the assessment of 30-day post-MI mortality and for the evaluation of care pathways (first pathway evaluated: chronic obstructive pulmonary disease), for which indicators will be calculated using data including medico-administrative health insurance data.

For the appropriate use of the indicators for other purposes (public disclosure, financial incentive to improve quality, etc.) additional validation is required (see Chapter IX).

3. Synopsis of the strategy implemented by HAS

The method implemented for the development, validation and use of this type of indicator is the fruit of HAS' capitalisation on the experience of countries which have been measuring outcome indicators for numerous years (Le Pogam et al. CLARTE, 2012), including in terms of vigilance as to their limitations (HAS, 2017; West et al. 2008; Quan et al. 2013). HAS has also based itself on international criteria for the validation of outcome indicators such as PSIs (Januel, DREES, 2011; Le Pogam CLARTE, 2012; Romano, 2009; Davies et al. AHRQ, 2001; Taffé et al. 2012), and on its experience in the development, validation, and management of national campaigns for care quality and safety indicators measurement.

Eighteen months are required for the development and validation of this type of indicator. To that effect, HAS relies on various players[1] and implements 8 equally important phases (see Figure 1).

It comprises an analysis of published literature, faced to the experts opinion of the working group and backed by experimentation:

- the literature review looks into the indicator's interest and definition (events, target population, risk factors);
- the experimentation comprises:
  - the translation of this definition into descriptive analyses based on data from available national databases (the PMSI's in-hospital data and National Health Insurance data (SNIIRAM)). These analyses contribute to the decisions to be made for the substantiated definition of the indicator,
  - a testing by patient record reviews by volunteer healthcare organisations: this fundamental step makes it possible to confirm that the indicator has the required reliability for a first data output aimed at the internal management of the quality of care and risks.

This strategy – from development to testing by patient record reviews (see Figure 1: stages 1, 2, 3 and 4) – allows the validation of outcome quality and safety indicators based on the following international criteria (see Table 2):

- clinical relevance of the indicator and construct validity;
- content validity (nosological framework);
- criteria validity and predictive validity;
- adjustment;
- discriminant validity:
  - inter-healthcare organisation variability,
  - difference actual versus a performance target,
- relevance for improvement;
- stability over time;
- potential negative side-effects.

1 The HAS Board, the project managers, the stakeholders group, the work group (made up of experts and service users), and the partnership with ATIH.

At stage 4, the indicator is validated for the internal management of the quality and safety of care.

The next stage corresponds to automatic indicator production and output to all healthcare organisations concerned (see Figure 1: stage 5).

Stage 6 is the adoption of the indicator objective, method and results by healthcare professionals in terms of clinical relevance, interpretation, as well as expected event-coding quality (applicable ATIH guidelines). This fundamental stage firstly provides feedback on non-anticipated situations via the literature review and working group, thus improving the reliability of the indicator (see Figure 1: stage 6). It also fosters knowledge and understanding of the indicator (objective, method, and results) by healthcare professionals in all concerned healthcare organisations:

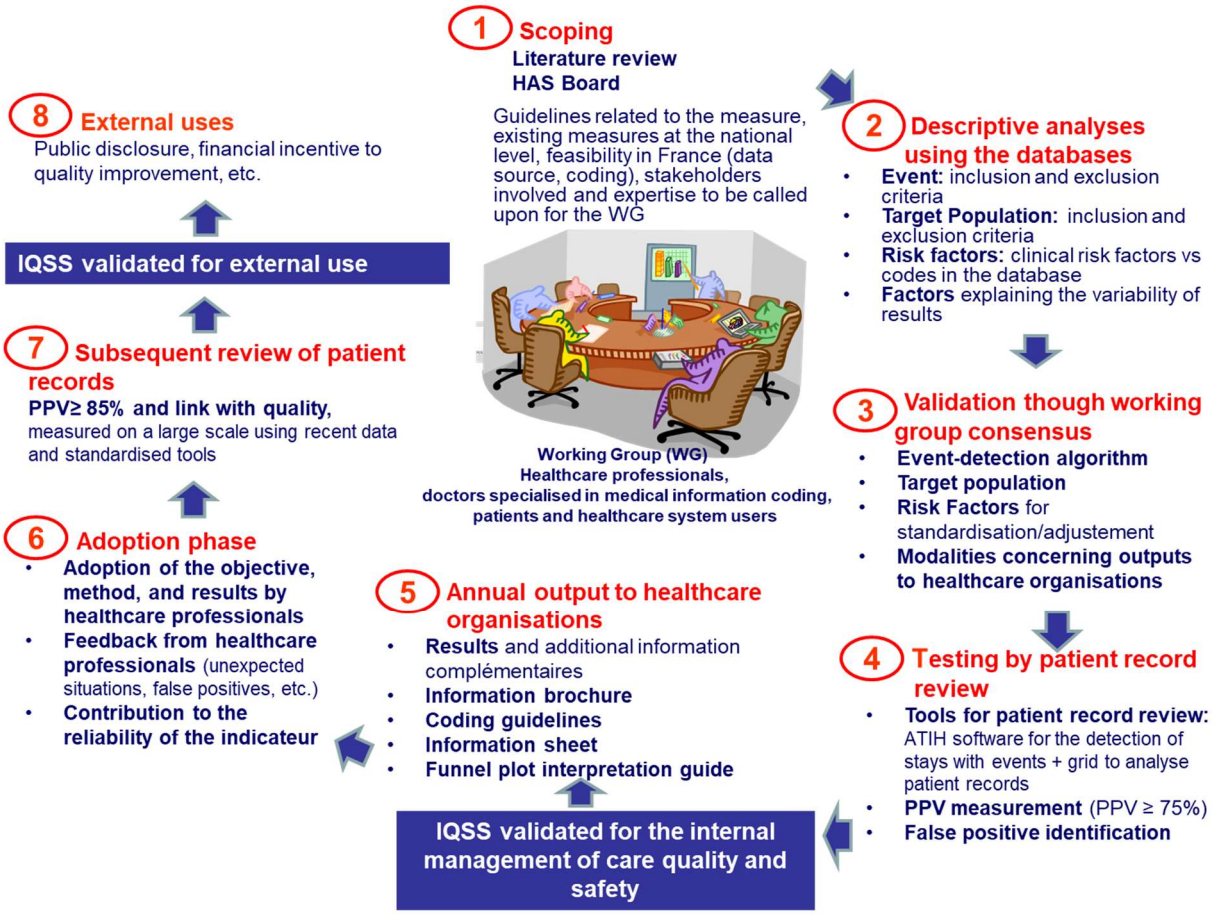
- this adoption phase allows the compilation of the healthcare organisations' first feedback on the identification of clinical situations or specific organisational aspects and false-positives detected (stays wrongly detected), for the potential optimisation of the indicator;
- the analyse and/or actions around this indicator can be used in the certification programme.

An additional validation stage is required for all external use. Patient record reviews at a distance from the first output are carried out, using standardised tools made available for the detection of stays with events and the analysis of the corresponding cases via the grid. It allows, after the adoption phase, the use of recent data for the large-scale measurement of the positive predictive value (PPV ≥ 85%) and the link with quality of care (see Figure 1: Stage 7).

Following this stage, external uses of the indicator may be envisaged in accordance with modalities to be defined (see Figure 1: Stage 8):

- public disclosure;
- financial incentive to improve quality.

Figure 1. Strategy for the development, validation and use of HAS' outcome quality and safety indicators: example of the measurement of a low-frequency event.



For each outcome quality and safety indicator produced by HAS, a development report is published. A descriptive analysis report on the national results is produced at each indicator output. The indicator is updated and optimised as required, in keeping with literature developments, improvements in the quality of coding and/or feedback from the healthcare organisations using it. The reports and tools are available online on the page dedicated to each indicator on the HAS website.

4. Players involved

4.1. The HAS Board

The HAS Board validates the scope of the indicator. The scoping process specifies the interest and stakes of the set-up of the indicator, the players involved in the provision of care, the expertise required to carry out the work, any partnerships that may be required, the expected time frame, information on the feasibility of the detection of the events sought in the target population, and the main references available.

4.2. Dual-expertise management team

The work is managed by a dual-expertise management team, made up of a scientist and a statistician.

4.3. Agence Technique de l'information sur l'Hospitalisation (ATIH)

ATIH is a partner in this work. It is notably involved in the validation of the codes used, the calculation of the results and complementary information once the indicator has been validated, the production of dedicated coding guidelines and softwares for the detection of stays in the PMSI, the management of the testing platform and the output of indicators and tools.

4.4. A collaborative approach

A collaborative approach is used. It brings together the players involved in the care whose indicator is the result. These players are identified and divided into 2 groups according to their role in the work: the stakeholders group and the experts working group.

Stakeholders group

A stakeholders group, composed of professional organisations and associations of patients/healthcare users concerned by the indicator is put in place. This group is informed by HAS of the progress of the work and, subsequently, of the national results and their trends after each output. The group may provide advance warning of a potential difficulty and promote the dissemination of the work and its results, in particular through publications and presentations at congresses.

Experts working group

A multidisciplinary working group made up of experts is put in place. It brings together the expertise of the medical staff caring for patients, doctors with expertise in the coding of medical information, patients and healthcare system users. The experts in the working group are approached through calls for candidates on the HAS website, via their professional organisations and associations and/or co-opted from other HAS working groups. The expert candidates' public declarations of interest (DPIs) and the information published on the government website – transparence.gouv.fr – are analysed by HAS' DPI validation committee. The experts working group is called upon for the scientific and clinical validation of the indicator, notably its criteria for the inclusion/exclusion of the target population and risk factors, and the analysis of the results after their output.

5. Development for the internal management of quality

The development is based on the scientific, clinical and statistical validation of the indicator: target population, event, risk factors and expression of the indicator.

5.1. Target-specific documentary search

A literature review focused on the event assessed in the target population is conducted. It covers published indicators measuring the event on the national scale, good practice guidelines and the expected target, risk factors associated with the event, and the factors explaining the variability of the results.

The working group's joint validation is sought concerning the scope of the measurement, the inclusion and exclusion criteria defining the target population, the numerator (event) and the denominator (when it is not the target population), the clinically relevant risk factors, and the codes available to identify with a relative reliability these elements in the database used.

5.2. Target population

Definition

The target population is composed of the stays or patients for which the event is detected.

The target population is systematically refined for each indicator. The aim is to target, through the specification of inclusion and exclusion criteria, a large, homogeneous population with room for improvement, in whom the occurrence of the event can be further reduced. This consists in identifying, based on existing literature, complex cases for which there is a major added risk not associated with the quality of care. Depending on their volume and/or impact on the occurrence of the event, these cases are either automatically excluded, or considered in the adjustment.

Examples of stays automatically excluded from the target population:

- for the in-hospital measurement of thromboembolic events: patient stays included in the database with coding error, transferred from another healthcare organisation or between healthcare organisations, palliative care, with the reason for admission (main diagnosis) being the event sought, discharge against medical advice or patients who ran away;
- for the 3-month measurement of surgical site infections or rehospitalisations 1 to 3 days after outpatient surgery: patients stays for which the data have not been linked or are incorrectly linked, patients residing abroad or for whom the place of residence is unknown, discharge against medical advice or patients who ran away, etc.

Descriptive analyses

The enumeration of the criteria to be analysed, as well as the measurement of their impact on the occurrence of the event, make it possible to select those to be excluded (low volume, population with a very high risk of morbidity/mortality, population cared for in specialised healthcare organisations) from those that will be included in the adjustment model (large volume, added-risk, population cared for in most healthcare organisations). The flowchart produced makes it possible to list and enumerate inclusions and exclusions to arrive at the indicator's target population (number and percentage of target stays/patients and healthcare organisations concerned).

5.3. Event

Definition

The identification of the events is based on an algorithm composed of the specific codes applicable, associated with non-specific codes if they are widely used in actual practice. For example, the event may be a complication (such is the case for standard patient safety indicators), a rehospitalisation or death.

Descriptive analyses

Events are counted in the database. For example, the descriptive analysis may concern the total number of events detected, the number of events detected over a particular follow-up period, or in a particular place (in or outside the healthcare organisation where the target population is cared for), etc.

5.4. Risk factors (adjustment or standardisation)

Definition

Risk factors are selected because they are clinically relevant, associated with an added risk of occurrence of the event not linked to the quality of care and identifiable in the database. Due to the fact that comorbidities are only coded if they use up resources during the target stay, their detection is done during the target stay and may be optimised by their detection in prior stays going back one or more years.

Descriptive analyses

Descriptive analyses will make it possible to measure the volume of each risk factor and its impact on the occurrence of the event. These are univariate analyses that make it possible to test the link between the risk factor and the occurrence of the event.

Adjustment model

The selected risk factors are introduced into the model. The selected factors are those which are significantly associated with the occurrence of the event and whose volume is non-negligible. However, a clinically validated risk factor with a non-negligible volume but no significant impact may initially be maintained in the model if an improvement of its coding can be envisaged through the provision of dedicated coding guidelines.

5.5. Defining the indicator

In the case of low-frequency events, the measurement sought is that of a standardised ratio of the observed over the expected number of events in the target population. The observed number is the number of events coded in the database. The expected number of events takes into account the factors associated with the risk, which are independent from the quality of care and are identifiable in the PMSI. It is calculated using a regression model covering the entire reference population (target stays/patients in the PMSI for the whole of France).

This ratio has the advantage of being compared to 1, a benchmark that does not vary from one year to the next, unlike the national event rate (which is not a clinically relevant target). Ideally, an event ratio must be compared to a clinically valid target (which is rarely available in published literature).

To minimise the risk of error, the result is compared to a benchmark, with a risk of error of 5% or 0.2%, depending on whether you use the limit of 2 standard deviations (2SD) or 3 standard deviations (3SD). For any use outside the healthcare organisation, HAS recommends the use of the 3SD limit, because of the minimal error risk. A result which is significantly higher or lower than the benchmark corresponds to a high outlier or a low outlier status, respectively. A result which is not statistically different from the benchmark corresponds to a status within the norm. For low-frequency complication-type events, the high outlier status corresponds to the worst outcome.

5.6. Testing by patient record review

This stage is fundamental for the validation of an indicator. It verifies the degree of reliability of the algorithm in detecting the event sought in the target population and ensures its validity for a first output intended for the internal management of the quality and safety of care.

The testing of the indicator by patient record review may be done by HAS or within the framework of a partnership with a research team, in accordance with the specifications laid down by HAS. The modalities of this testing are described in Table 2 (criteria validity).

It consists in asking volunteer healthcare organisations to analyse patient records for the corresponding stays/patients detected by the algorithm tested in the database. In the case of stays or hospitalised patients, a healthcare organisation can only access the records of patients hospitalised in its own facilities. This analysis is conducted in compliance with CNIL data confidentiality requirements, in particular concerning access to patient records and the confidentiality of the data collected. The patient record review is ideally done for all of the stays detected, or if applicable, for consecutive stays or randomly selected stays.

For this stage, HAS provides institutions with the following on a secure platform:

- software for the detection of stays with events in their database (example: PMSI) produced by ATIH;
- a grid for the collection of the information required for the validation of the indicator.

HAS analyses all of the information uploaded by participating healthcare organisations.

For rare events – such as in-hospital coded thromboembolic events in adult patients undergoing total hip or total knee arthroplasty (ETE-ORTHO) – the indicator's PPV is calculated (Positive Predictive Value: percentage of stays with events detected in the database, confirmed in the corresponding patient records), and any false positives are identified (situations wrongly identified as events and/or target stays). Whenever possible, false positives are excluded beforehand in order to improve the reliability of the indicator in detecting the events sought in the target population.

A predictive value of at least 75% is required for a first output of results to the healthcare organisations, to be used internally for the improvement of practices (Le Pogam et al. CLARTE, 2012, Januel, DREES, 2011, Hefner et al. 2017).

6. Criteria for the validation of outcome quality and safety indicators for internal management purposes

Following the development phase and the experimentation phase (including the testing by patient record review), the indicator is validated based on international criteria (Januel, DREES, 2011; Le Pogam CLARTE, 2012; Romano, 2009; Davies et al. AHRQ, 2001; Taffé et al. 2012):

- clinical relevance of the indicator and construct validity;
- content validity (nosological framework);
- criteria validity and predictive validity;
- adjustment;
- discriminant validity:
  - inter-healthcare organisation variability,
  - deviation from a performance target,
  - relevance for improvement;
- stability over time;
- potential negative side-effects.

Their definitions and evaluation modalities are presented in Table 2 appended to this document.

7. General dissemination for the internal management of quality

The indicators validated by HAS between 2015 and 2018 assess the quality of care and safety of hospitalised patients. These indicators are produced and disseminated automatically to all relevant healthcare organisations via a secure platform, and are supplemented with tools, documents and additional information for the analysis of results.

The HAS modalities used for the output of the results provided for improvement purposes are described below.

HAS also provides ongoing assistance, answering the healthcare organisations' questions via the Agora online platform and by email at contact.iqss@has-sante.fr (during the testing by patient record review or the output of the results).

Quality-improvement tools

To help improve the quality of care, practices and risk management, each indicator is backed by:

- the current professional/organisational/regulatory guidelines related to the event measured;
- risk management tools;
- and/or any other HAS tool to improve patientcare and safety, communications, team work, etc.

Information sheet

An information sheet is produced for the validated indicator, detailing the following: the importance of the topic, with a reminder of the professional references associated with the measurement, the definition of the measurement, the objective of the measurement, the type of indicator (outcome), the data source, the validation, the data collection method, the nature of the indicator (mode of expression), the target population, the event, the risk factors, the calculation, the modality used for the output of the indicator results, the pace of dissemination of the results to healthcare organisations, the additional information provided for the contextualised analysis of the indicator results, the factors explaining the variability of the results, the modalities for using the indicator and the main references. This data sheet is published on the HAS website and is available to healthcare organisations, along with their results, on the QualHAS output platform.

Information brochure

An information brochure is produced to inform users of the measurement in healthcare organisations.

It provides answers to the following questions: Why this indicator? What does it measure? Who is concerned by this indicator in healthcare organisations? How is this indicator produced? How is this indicator to be used?

A memo is produced concerning care quality and safety indicators (IQSSs), their advantages and limitations when they are measured from medico-administrative databases, and the link to the page dedicated to the indicator on the HAS website.

Guidelines concerning the coding of events in the database

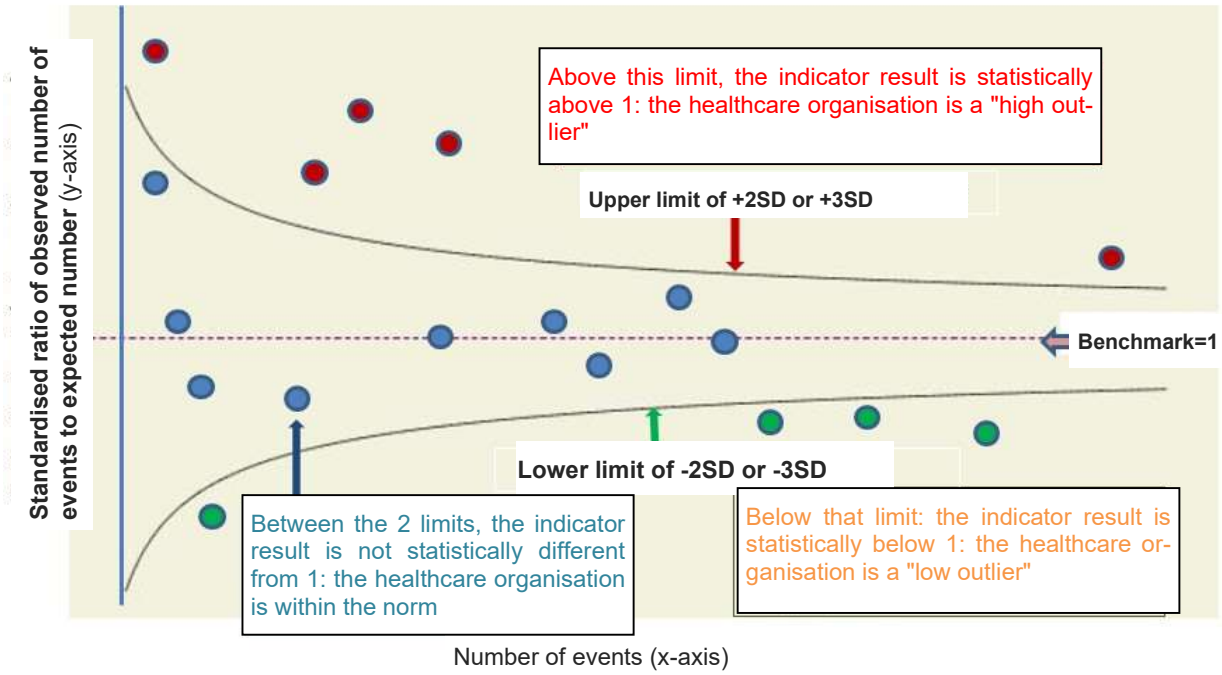
For the measurement of events such as complications, ATIH produces a reminder of the guidelines for the coding of these events in the PMSI. It is available in QualHAS with the indicator and is published on the dedicated page on the HAS website. These specific coding guidelines are intended to improve the quality of the coding of the events measured, and thereby optimise the reliability of the indicator.

Result output modalities

Outcome indicators measuring low-frequency events can be disseminated to each healthcare organisation in a funnel plot (Spiegelhalter a, 2005). This is an easy-to-read graphic representation, characterised by 4 parameters (see Figure 2):

- an indicator (on the y-axis): for example, the value of the standardised ratio of the observed number of events to their expected number;
- a target (the continuous horizontal line): the reference value is set at 1 if the indicator is an observed/expected ratio;
- a precision parameter (on the x-axis): the number of expected events, which is given preference over volume for low-frequency events;
- control limits (the funnel) for which the probability for an institution to be outside those limits is p. The probability of being wrongly statistically considered as different from the benchmark for healthcare organisation above or below the limit is 5% for the 2SD limit, and 0.2% for the 3SD limit.

Figure 2. Funnel plot



Each healthcare organisation sees its results in colour, along with the anonymised results of the other healthcare organisations in another uniform colour. Each result is to be compared to the benchmark, which is 1 for the observed/expected ratio. Reliance on the limit of +3SD in the funnel plot is given preference due to its low risk of error at 0.2%. However, for pedagogical purposes and to stimulate institutions, the output may place the result within the norm, between 2SD and 3SD and beyond 3SD.

The use of an over-dispersion[2] factor is to be examined on a case-by-case basis (Spiegelhalter b, 2005).

The causes of a result "outside the limits" are to be investigated by each healthcare organisation concerned, through a patient record review.

Minimum threshold of target stays: a minimum threshold of at least 10 target stays is generally used by HAS for national analyses. This threshold is arbitrary and is not based on any available scientific or statistical argument. HAS thus uses a threshold of 10 target stays for its indicators, compared to the threshold of 25 target stays used by the United States and the threshold of 50 target stays used by the OECD. HAS privileges to include a maximum number of institutions in the improvement approach with a minimum acceptable threshold set at 10 target stays. For score-type indicators (e.g. the quality of the hospital discharge letter), the threshold is set at 30 target stays.

Funnel plot interpretation guide

When the indicator is represented in a funnel plot, it is supplemented with a guide to facilitate the interpretation of the results. It explains how to read the results and what conclusions may or may not be drawn from the funnel plot (see Table 1).

Table 1. What the funnel plot can and cannot be used for.

The funnel plot allows:	The funnel plot does not allow:
To answer the question: is a healthcare organisation's result different from the target value? For each healthcare organisation to: – clearly see their results among those of the other healthcare organisations, – compare their results with the benchmark (target value) = 1, – position themselves inside or outside the funnel plot defined by the 2SD or 3SD limits, – have a non-standard or outlier status, if their results are outside the limits.	– Compare healthcare organisations. For each healthcare organisation, compare their results over time, as the result shown in the funnel plot depends on parameters measured using the data for the year under study. In addition to the annual funnel plot, results covering several years are provided to each healthcare organisation. – Explain the reasons for a "non-standard or outlier" status.

Additional information

Useful information for the analysis of the indicator results that can be calculated from the database is identified and automatically provided to the healthcare organisations along with the indicator results. This notably includes the number of target stays, coded events and risk factors calculated for the healthcare organisation's target population and at the national level. This information concerns the population targeted by the indicator, making it possible, in certain cases, to identify the care to be investigated by a patient record review.

2 We talk about over-dispersion when an excessive proportion of institutions are outside the funnel limits. This phenomenon is observed when the adjustment model is incomplete: such is the case when risk factors impacting the results are not measured or incorrectly measured. An over-dispersion factor can then be introduced in the calculation of the limits, which will alter their positioning and thus reduce the number of outlier institutions.

Results covering a period of at least 3 years are retrospectively provided to inform healthcare organisations of their result trend over time. This is not a comparison over time, as each output covers a different patient population.

To analyses the causes of the occurrence of an event, healthcare organisations can also use the information provided in the information sheet: guidelines for good clinical/organisational practices linked with the result measured, as well as the factors explaining the variability of the results identified in published literature, but which are not measurable in the database.

The HAS modalities of the indicators output allow their use for the internal care quality and risks management. The use of this type of indicators (outcome and practices analysis, review of patient records and/or implementation of improvement actions) are to be used in the certification programme right from the first output. HAS' certification has the ambition to encourage healthcare organisations with a "high outlier" status to analyse their results and identify and implement improvement actions.

8. Adoption of the indicator by healthcare professionals to improve its reliability

The first output of the indicator results and additional information marks the start of the phase of adoption by healthcare professionals.

This is a fundamental stage that:

- enables healthcare professionals to become acquainted with the indicator, grasp its complexity, its interest and its interpretation, become familiar with current coding guidelines, conduct a patient record review, and submit their questions and observations to HAS;
- provides HAS with feedback on actual field practices to supplement the data in published literature and working group appraisals – for example, the knowledge of specific clinical situations or organisational practices that could explain their outlying results or the false positives recorded (stays wrongly detected by the algorithm). This helps to improve the reliability of the indicator for the detection of the event sought in the target population.

9. Additional validation for external uses

9.1. Patient record review, at a distance from the first output

Before any external use, the predictive value should be measured through a patient record review, at a distance from the first output. This ensures that, based on recent data and after the probable improvement of coding quality the required predictive value is achieved. It also makes it possible to get field players involved and reassure them as to the reliability of the indicator. This subsequent record review is carried out in the same ways as for the initial validation test, using the same tools calibrated on the current version of the algorithm and the source data. The objective is to measure the predictive value and, where applicable, identify any false positives, analyse the potential causes of occurrence of the events and identify those related to the quality of care, and those which are potentially avoidable.

For example, for ETE-ORTHO, which measures a low-frequency event, a review of the patient records was conducted on the 3rd year of output. It made it possible to verify the reliability of the indicator by re-calculating the PPV using recent data and envisage a use other than the internal management of quality.

The predictive value conditions for external use are valid for outcome indicators that measure specific events (output) such as post-operative complications. When results concern non-specific multi-factor events, such as rehospitalisations or death (outcomes), the predictive value is not sufficient on its own to envisage external use: the link between a high outlier result and quality of care must be ensured with a minimal risk of error. The closer the event to the care provided, the more probable the link. The patient record review then provides the advantage of mobilising field players to quantify the events potentially linked to a medical cause and, out of these causes, identify those that are potentially avoidable and for which a corrective action is possible. When such potentially avoidable quality-related events make up the majority of events, the indicator becomes "discriminatory and relevant for improvement purposes" and external use can be envisaged.

9.2. Modalities concerning external uses

After an adoption phase of 1 to 2 years, and if the predictive value is at least 85%, uses other than the internal management of quality may be envisaged: publication, financial incentive to improve quality, etc.

Given the coding limitations linked to the data source for this type of indicator, the only certainty concerns healthcare organisations with high outlier results: their observed number of events is significantly higher than the expected number with a minimal risk of error (equal to 0.2%). The results of low outlier healthcare organisations are potentially better than others but without any certainty as to the assumption that this may be due to better-quality practices.

Thus, for external uses, high outlier healthcare organisations are considered as grade C, by analogy with IQSS ratings, while the others are considered within the norm, i.e. as grade A or B.

Public disclosure

For example, the ETE-ORTHO indicator has been disseminated on the Scope Santé website. High outlier healthcare organisations appear in orange, which means "lower than expected outcomes", while the others appear in green which means "outcomes within the norm or potentially better than expected".

Financial incentive to improve quality

For example, the integration of the ETE-ORTHO indicator in the 2018 quality improvement funding scheme consists in applying a moderate penalty to the IFAQ scores of high outlier healthcare organisations. At present, the use of this type of indicator in new funding models has not yet been defined.

Bibliographic references

Agence Nationale d'Accréditation et d'Évaluation en Santé. Construction et utilisation des indicateurs dans le domaine de la santé: principes généraux. 2002.

Shaw, C, Kalo, I (2002). A background for national quality policies in health systems. Copenhagen, WHO Regional Office for Europe. http://www.euro.who.int/__data/assets/pdf_file/0019/142066/E77983.pdf Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260(12): 1743-8.

Direction de la recherche des études de l'évaluation et des statistiques, Januel JM. Développement d'Indicateurs de la sécurité des soins (PSI) à partir des bases de données médico-administratives hospitalières. Série sources et méthodes n°20. Paris: DRESS; 2011. http://drees.social-sante.gouv.fr/IMG/pdf/seriesource_method20.pdf

Le Pogam MA, Januel JM, Colin C, Consortium CLARTE. Modalités d'utilisation et de diffusion des PSI

(Indicateurs de sécurité des soins hospitaliers) dans les pays de l'OCDE. Saint-Denis La Plaine: HAS; 2012. https://www.has-sante.fr/portail/jcms/c_1262821/modalites-dutilisation-et-de-diffusion-des-psiindicateurs-de-securite-des-soins-hospitaliers-dans-les-pays-de-locde-2012?xtmc=&xtcr=1

Heath I, Hippisley-Cox J, Smeeth L. Measuring performance and missing the point? BMJ 2007;335(7629):1075-6.

Gubb J. Have targets done more harm than good in the English NHS? Yes. BMJ 2009;338:a3130.

Haute Autorité de Santé. Indicateurs de mortalité hospitalière : expériences étrangères, enseignements de la littérature et recommandations pour l'aide à la décision publique et le développement d'indicateurs en France. Indicateurs pour l'amélioration de la qualité et de la sécurité des soins. Saint-Denis La Plaine:

HAS; 2017.https://www.has-sante.fr/portail/upload/docs/application/pdf/2017-

10/rapport_mortalite_2017.pdf

Steven A. Farmer, MD, PhD; Bernard Black, MA, JD; Robert O. Bonow, MD, MS. Tension Between Quality Measurement, Public Quality Reporting, and Pay for Performance. JAMA. 2013;309(4):349-350. doi:10.1001/jama.2012.191276.

Farmer SA, Black B, Bonow RO. Tension between quality measurement, public quality reporting, and pay for performance. JAMA 2013;309(4):349-50.

Januel JM, Couris CM, Luthi JC, Halfon P, Trombert-Paviot B, Quan H, et al. Adaptation au codage CIM10 de 15 indicateurs de la sécurite des patients proposés par l'Agence étasunienne pour la recherche et la qualité des soins de santé (AHRQ). Rev Epidémiol Santé Publique 2011;59(5):341-50.

West AN, Weeks WB, Bagian JP. Rare adverse medical events in VA inpatient care: reliability limits to using patient safety indicators as performance measures. Health Serv Res 2008;43(1 Pt 1):249-66.

Quan H, Eastwood C, Cunningham CT, Liu M, Flemons W, De Coster C, et al. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). BMJ Open 2013;3(10):e003716.

Direction de la recherche des études de l'évaluation et des statistiques, Haute Autorité de Santé, Januel JM. Les méthodes d'ajustement dans les modèles d'évaluation de la mortalité hospitalière. Partie 1 – Étude descriptive. Série études et recherche n°112. Paris: DRESS; 2011.

Analyse de la Mortalité Post-Hospitalisation en France en 2008-2010 (AMPHI) : recherche d'Indicateurs par établissement- Rapport final. Lamarche-Vadel A, Ngantcha M, Jougla E, Rey G. CépiDc-Inserm, January 2014. Report submitted to the Haute Autorité de Santé (HAS), DREES and CNAMTS.

Spiegelhalter DJ.a Handling over-dispersion of performance indicators. Qual Saf Health Care 2005;14(5):347-51.

Spiegelhalter DJ.b Funnel plots for comparing institutional performance. Stat Med 2005;24(8):1185-202.

Patrick S. Romano, MD MPH. Updated validation of AHRQ Patient. Safety Indicators in the USA. UC

Davis Center for Healthcare Policy and Research. Organization for Economic Cooperation and Development. 2009

Davies SM, Geppert J, McClellan M, et al. Refinement of the HCUP Quality Indicators. Technical Reviews, No. 4. Rockville (MD): Agency for Healthcare Research and Quality (US); 2001 May

Patrick Taffe´, Patricia Halfon, William A. Ghali, Bernard Burnand, for the International Methodology Consortium for Coded Health Information (IMECCHI)Test Result–Based Sampling: An Efficient Design for Estimating the Accuracy of Patient Safety Indicators. Med Decis Making 2012;32:E1–E12.

Jennifer L Hefner, Timothy R Huerta, Ann Scheck McAlearney, Barbara Barash,Tina Latimer, and Susan D Moffatt-Bruce. Navigating a ship with a broken compass: evaluating standard algorithms to measure patient safety. Journal of the American Medical Informatics Association, 24(2), 2017, 310–315.

Haute Autorité de Santé. Validité des indicateurs de résultats mesurés automatiquement à partir des bases de données médico-administrative. Note technique, HAS 2019.

Appendix. Criteria for the validation of HAS' care quality and safety indicators

Table 2. Criteria for the validation of an outcome indicator measured from medico-administrative databases

Parameters assessed	Definition/Fields Concerned	Assessment method
Clinical relevance of the indicator and construct validity	Clinical relevance is the indicator's potential link with the quality and safety of care. Construct validity is the ability for the indicator to be correlated to other indicators assessing patient care (e.g. indicators concerning recommended clinical practices and/or patient outcomes).	**Qualitative** Consensus of the work group (WG) Literature: e.g. existence of professional guidelines for good practices in connection with the event measured by the indicator. **Quantitative** Ad hoc investigation through a patient record review: e.g. for all-cause rehospitalisation or death indicators, the link with the quality of care must be confirmed for the majority of stays/patients detected with events.
Content validity (nosological framework)	Indicator's ability to represent the major dimensions of a concept of interest. It concerns the translation of the clinical definition of the events sought (numerator) in the target population into codes/variables available in the database.	**Qualitative** Consensus of the GT The target population is defined by targeting the majority of patients cared for in the institutions concerned, for which a degree of improvement is expected, after exclusion of situations involving added risks not linked to the quality of care. These codes/variables are updated at each change/modification in nomenclature and/or coding guidelines, in coordination with ATIH.

Criteria validity and predictive validity	Criteria validity measures the indicator's ability to produce results which are comparable to those produced using the "gold standard". It concerns the ability to identify the events sought in the target population. Predictive validity is the ability to produce the result of interest. For low-frequency events, HAS uses the PPV. For high-frequency events the NPV is opted for (see Note technique, HAS 2019 in the references and online). ➔ **The positive predictive value** (PPV) corresponds to the probability that an event has actually occurred if the tested algorithm detected it. This is the number of true positives divided by the total number of events detected. $VPP = \frac{A}{(A+B)} = Proba(GS_+ \mid BDMA_+)$ ➔ **The negative predictive value** (NPV) corresponds to the probability that an event has not actually occurred if the tested algorithm has not detected it. This is the number of true negatives divided by the total number of cases with no events detected. $VPN = \frac{D}{(C+D)} = Proba(GS_- \mid BDMA_-)$ ➔ **False positives:** These are events and/or target populations that have been wrongly identified via the database, and are not confirmed in patient records. They correspond to situations that, insofar as possible, should be automatically excluded in order to improve the indicator's PPV.	**Qualitative** Feedback from doctors specialising in medical information in the WG and ATIH on the reliability and actual use of the codes/variables available in the database. **Quantitative** The measurement of the PPV seems relevant to validate an indicator measuring a low-prevalence event, while the measurement of the NPV will be used to validate an indicator measuring a high-prevalence event. For HAS' outcome quality and safety indicators concerning low- or high-frequency events: HAS targets a number of 500 records and considers the representativeness of the institutions (category and status – outlier or within the norm) and events. This number of records would ensure the reliability of the measurement (within a margin of 5%), provided that it reaches at least 92.5% The predictive value must be at least 75% for a general rollout aimed at improving practices, and at least 85% for any other use (public disclosure, pay-for-performance programme, etc.).
Adjustment - Standardisation	Indicator's ability to take account of variables influencing the results independently from the quality of care (e.g. age, gender, comorbidities, etc.). This minimises the bias.	**Qualitative** Available literature and GT consensus, reliability of the identification of codes/variables in the database corresponding to standardisation/adjustment factors.

	The category of institution or the territory/region may be variables that explain the variability of results between institutions, but not adjustment factors. HAS considers that the same level of quality and safety are required irrespective of the category of institution or its region/territory.	**Quantitative** In the case of a regression model, the performance of the model is assessed according to: – the model's discriminant power measured through the calculation of the area under the ROC curve (C-stat); a value ≥0.70 is considered as satisfactory; – the model's good fit for the data evaluated using the Hosmer-Lemeshow test: a non-significance of the test (p>0.05) has been considered as satisfactory; – the calculation of the Akaike information criterion (AIC) and Bayesian information criterion (BIC): the model with the lowest value for those two criteria is considered as the best suited to the data. Possible interactions between the models' variables are systematically sought.
Discriminant validity	The indicator's ability to measure variability between institutions and set an improvement target in relation to a benchmark.	
➔ Variability between healthcare organisations	Indicator's ability to discriminate between healthcare organisations by observing the variability of their results.	**Quantitative** Example: dispersion of measurements between healthcare organisations in relation to a benchmark in a funnel plot. Results are calculated for each healthcare organisation in relation to the 2SD limit (5% error risk) and 3DS limit (0.2% error risk). The 3SD limit is preferable to identify healthcare organisations which are significantly different from the benchmark, especially in the case of external use. Over-dispersion factor: it can be used when the adjustment does not take account of factors affecting the event measured.

		The percentage of outlier healthcare organisations should be around 10% for the implementation of this mechanism.
➔ Deviation from a performance target	The indicator's ability to identify scope for improvement through the observation of a deviation from a performance target (example: a benchmark published in a literature review or a national benchmark).	**Quantitative** Performance threshold to be defined. Example: Number/rate of observed events exceeding the expected number or a published clinical target.

➔ Relevance for improvement	Ability to act to improve results.	**Qualitative** Example: identifying the reason for the occurrence of an event deemed avoidable, for which an improvement action is possible. The improvement action is reflected in a more or less rapid improvement of outcomes (e.g. the reduction of mortality).
Stability over time	Tool's ability to produce consistent results over time.	**Quantitative** Measurement/Follow-up of the national stability observed over a period at least 2 years before the implementation of the measure. Monitoring of the stability of the measure over time is required and any sudden change must be analysed with the health professionals/coders involved.
Potential negative side-effects	The indicator's potential to induce a behaviour that will modify the results without any connection with the improvement of the quality of care (gaming).	**Qualitative**. Analysis of publications on adverse effects when the measure is used for purposes other than the internal management of quality, e.g. for publication purposes and/or quality-based financial control. **Quantitative**: follow-up of changes in the indicator, along with the analysis of parameters that can improve the indicator results without affecting the quality of care (e.g. reducing the duration of stays, discharges through transfers, etc.).

Retrouvez tous nos travaux sur

www.has-sante.fr

HAUTE AUTORITÉ DE SANTÉ

Développer la qualité dans le champ sanitaire, social et médico-social