

Moving data collection from a “vicious” cycle to a “virtuous” cycle – a cycle of continuous improvement.

Authors

Paul O’Connor, Team Lead*

Beth Reid, Senior Consultant*

Lee Ridoutt, Principal Consultant*

Cliona O’Donovan, Quality Assurance & Operations**

Bin Jalaludin (Jal), Senior Consultant*

Ric Marshall, Senior Consultant*

*Digital Health Unit *Human Capital Alliance Pty. Ltd.*

**National Office of Clinical Audit, Ireland

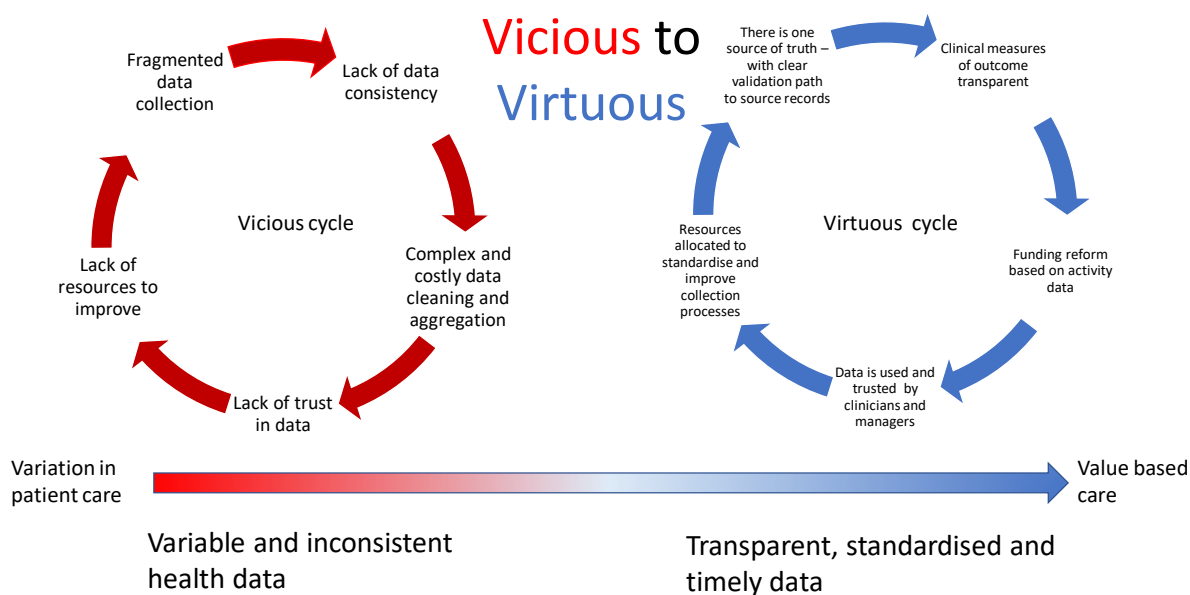
Abstract

Improving the quality and usefulness of the data we collect from the health care system is a constant challenge.

Health care data collection can be locked into a “vicious” cycle, where ultimately the data are not used because the potential users, such as clinicians and managers, think that the quality of the data is not sufficient to support good decision making. The clinicians, clinical coders, managers, and data custodians who are responsible for originating and collecting the data (together the data collectors) cannot see that their effort is contributing to improving decision making and enhancing patient care and public health. How can these data collectors maintain their commitment to producing high quality data if it is not used? Our experience is that this lack of engagement leads to further fall in the quality of the data, producing poor, unusable or unused data.

Variable and inconsistent data cannot support value-based care. In this paper we postulate how data collectors can best turn a vicious into a virtuous cycle.

Figure 1 Vicious cycle to Virtuous cycle



A virtuous cycle of quality improvement occurs when the data is trusted, which leads to the appropriate resources being allocated to standardise and improve the data collection process resulting in “one source of truth” with clear validation paths to the source documents.

There are a number of internationally recognized data quality frameworks used to measure relative data quality. Our work over the last six years in Australia, Republic of Ireland, Singapore, Kingdom of Saudi Arabia, and Fiji. takes these frameworks and focuses on the specific factors that affect quality and motivation of the data collectors operating in acute care hospitals. Our method was to undertake detailed reviews of original data sources, assessment of the quality of the collected data (over 30 million acute patient journeys), benchmarking processes against best practice, and hundreds of structured and semi- structured interviews and workshops with data stewards, custodians and key stakeholders. This deep pool of insights was then analysed and discussed by those on our author’s panel each coming with a different set of experiences and perspectives leading to the insights we share in this paper.

In this paper, we discuss the factors that influence and support data gatherers in high quality data collections: these are grouped under the categories of data, governance arrangements, collection processes, infrastructure and people - the collectors and the users. All factors are not equal in their influence. In particular in our view standardising data to enable aggregation is a fundamental foundational element. So too are whether the information technology systems can facilitate the collection and aggregation of data. Training, rewards, clear and available career pathways, regular auditing, cross checking of the validity of the data from other sources and promoting the use of data with appropriate feedback loops are also important.

Our conclusion is that the timely use of trusted data by clinicians and other decision makers as close to the data source as possible is key. Making this data important and transparent in supporting funding and clinical care decision will ensure resources and recognition are provided to the data gatherers, leading to a virtuous cycle of quality improvement.

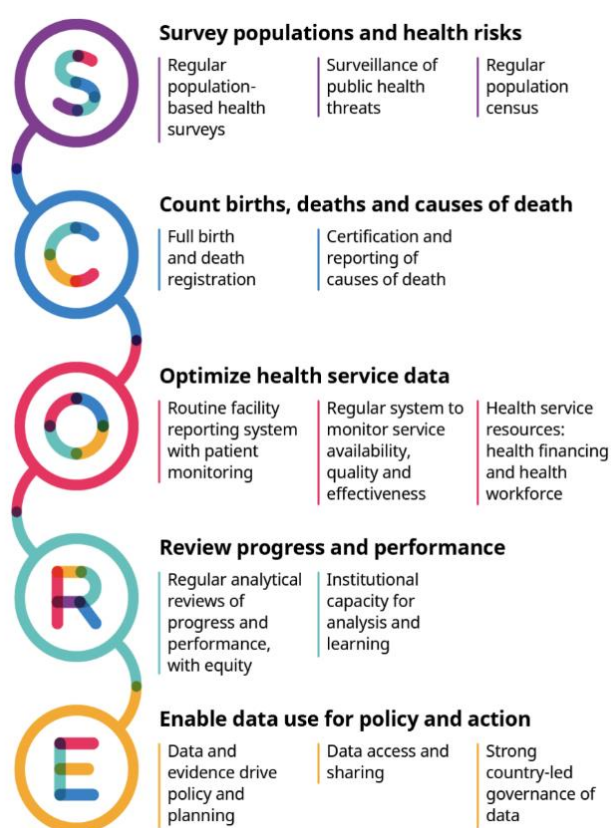
Background

Regular, reliable data from health facilities and the resource systems that support them is central to ensuring the availability and quality of health services.

Sound data that properly reflects the activity in the sector is essential to understanding what is happening to patients and how those responsible are responding to their needs. Health care is complex and its data collection systems and infrastructure are expensive. Countries around the world have invested in data collection approaches to support decisions taken as to health services, however, in many cases real value from such investment has not been realised.

¹The COVID-19 pandemic has created an unprecedented demand for high-quality health data and also highlighted many long-standing gaps. Health information systems around the world today are being stretched like never before. Not only do these systems need to track COVID-19 for effective pandemic response and recovery, but they must also continue tracking and responding to other health priorities.

Figure 2 SCORE Global Report 2020 - WHO



SCORE developed by WHO was used to present findings from the first ever global assessment of country data and health information systems, covering 133 country health information systems or 87% of the world's population. Key insights;

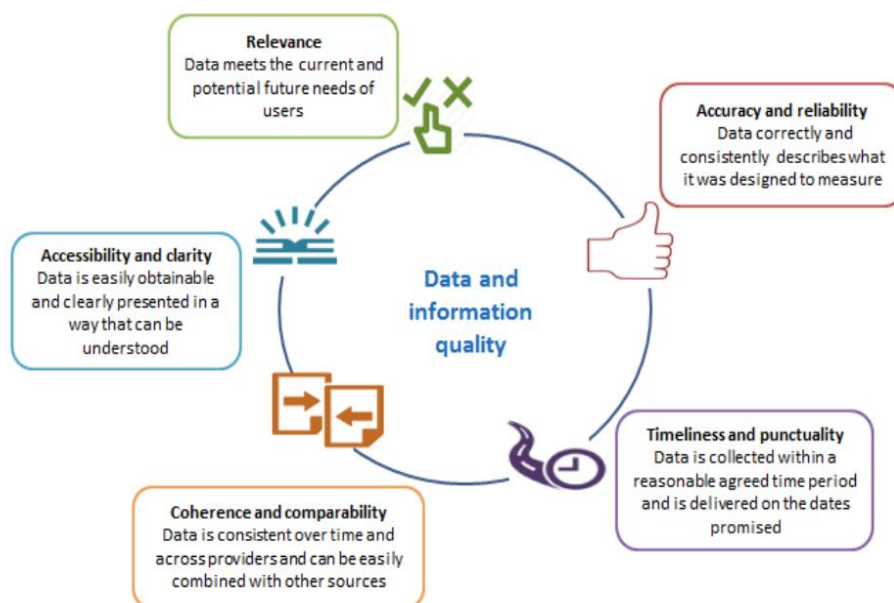
- Optimizing health service data is critical to ensure equitable, quality services for all. Yet 50% of countries have limited or less capacity for systematic monitoring of quality of care.
- 59% of countries have well-developed or sustainable capacity for using data to drive policy and planning, but only 42% of countries have good capacity for data access and sharing.
- Optimising Health service data based on International Classification of Diseases is the key focus area of our study

The dimensions of data quality used in one Data Quality Framework² suggest that a “target for health and social care organisations is to produce data that is sufficiently accurate, timely and consistent to make appropriate and reliable decisions, rather than aiming to produce perfect data.”

High quality data are:

- used to deliver and monitor safe and high-quality care for everyone
- of the highest quality and, where appropriate, collected as close as possible to the point of care
- data which are collected once and used many times
- ‘fit for purpose’ and cost-effective.

Figure 3 Data quality dimension Guidance on a data quality framework for health and social care Health Information and Quality Authority Ireland



The data to information life cycle consists of five steps: capture, submit, process, analyse, and disseminate. Problems can occur at every step of this cycle:³

- **Data gathering and integration.** Possible problems include fallible manual entry, biases (e.g., upcoding in claims), erroneous joining of tables in an electronic health record and replacing missing values with default ones.
- **Data storage and knowledge sharing.** Potential problems are lack of documentation of the data model and lack of meta-data.
- **Data analysis.** Problems can include incorrect data transformations, incorrect data interpretation, and use of inappropriate methodology.
- **Data publishing.** Problems include that the publication is not timely or relevant.

Data quality must be consistently defined in context⁴. Data quality assessment frameworks exist, however there is a lack of commonly agreed terminology and methodologies to assess quality or data maturity. Improving current practice requires understanding the maturity of current processes and what constitutes best practice.

It is important that an assessment of data quality as well as the actual and potential impacts of using the data are to be viewed together. For example, a reported high quality outcome measure may be the result of poor-quality reporting of complications rather than good clinical performance. An understanding of the data quality is especially important in establishing predictive modelling.

There are several standard data quality assessment frameworks and communities⁵ that offer open-source methods of measurement of data quality across the data life cycle⁶. Adopting a uniform method allows benchmarking and transparency across jurisdictions and classifications. Data capture and classification are important foundations for value-based care, and require an internationally recognised standard method to measure data quality or maturity.

“Trust and willingness to share data for public good and scientific advancement is a core requirement. Data quality assessment, management, visualization, and sharing requires an optimal

balance of privacy and security arrangements with adherence to FAIR principles. An ethical and secure framework based on public good is essential to produce a data asset that is fit for purpose.

Comprehensive Data quality assessment requires a culture of reciprocity, transparency, and interoperability across the data production and curation life cycle. Effective data quality assessment is underpinned by rigorous documentation at point of care, good management, and appropriate governance across the data production and curation life cycle.”⁷

It is possible to waste resources chasing high quality data. We must start with the data that can be collected currently, and then develop and apply standard measures of data quality.

Much can be done to improve the quality and usefulness of the data we collect from the health care system. Continuous data quality improvement is necessary to ensure that the data are trusted by clinicians and managers. Using trusted data improves decision making for value-based care.

Methods

The methods used in our work over the last 6 years were employed for the first time in Ireland⁸ and subsequently refined with the Ministry of Health (MOH) in Singapore and MOH in Kingdom of Saudi Arabia. The method represents a comprehensive review of data quality and the people and processes underpinning the clinically coded data in acute care hospitals. The work focused on establishing a measurement of process and outcome data against an established best practice in data collection and classification as a foundation for funding and patient care.

Specifically, a combination of data analysis tools⁹ that interrogated the full data set for completeness, specificity, sequencing and expected outcomes against international best practice and standards was conducted. In a number of situations extensive data cleaning was necessary to aggregate data to a level of analysis. This data analysis was triangulated with physical audits of representative¹⁰ samples of source documents and our learnings from structured interviews to establish a benchmark against a derived best practice.

Insights were produced for all health facilities, regional and national health authorities to test the insights and obtain feedback. These insights were further refined through hundreds of workshops with clinicians and the collectors of the data in developed and developing health sectors. Details of the workshop findings and interviews for Ireland have been published previously.¹¹

The insights gained from this analysis were tested and discussed with the author¹² panel who each bring unique and extensive experience to reach consensus on the overriding themes which emerged.

There are a number of Irish specific quality frameworks cited with examples of best practice. These were sourced from the National Office of Clinical Audit, Ireland and were not part of our formal studies but provide good frameworks of processes that have been based on international best practice.

Findings

We found that the best quality data existed where there was a high level of interest in using the data, where the data were used for meaningful purposes as close to the point of data collection as possible, and where the data was shared widely.

Thus, in our view, the key drivers of ongoing improvements in data quality are:

- **Interest** in using coded data by the clinicians and managers for:
 - Measuring the quality of care

- Outcome measures e.g., Hospital-Acquired Complications, routine clinical audits
- Funding/invoicing
- Understanding cost of care
- Understanding efficiency of care
- **Transparency** – the more data is shared the better the quality

There are a number of different factors which influence the data collection process including adequate resourcing, governance, people and infrastructure. However, we have found that if there is no interest in the use of the data by clinicians and managers, data quality (consistency, timeliness, and accuracy) improvements are difficult. Often this lack of interest in the data has been driven by a lack of understanding by those managing the data collectors as to how the data could be used to improve value-based care. Our work identified many examples of data quality “evangelists” who advocated to the key stakeholders the power of the data to transform patient care.

Health system data are collected for many purposes, including as the basis for funding and payment systems, comparison of clinician and hospital/facility performance, consumer feedback and public health decision making. All these purposes can contribute to improving the health system, but some will have a stronger influence on the motivation of the data gatherers to ensure the accuracy of their work.

The impact of the use of the data on its quality varies with the perceived importance of how the data are used. When we asked coders and coding managers about uses of the coded data, a frequent answer was “for statistics”. We probed deeper because we wanted to assess the interviewee’s understanding of the type and purposes of the statistics. Answers ranged from a good understanding of the statistics that were useful to the hospital to “just statistics” such as the top 10 diagnoses. A poor understanding of how the data are used will likely have a negative impact on the motivation of the coders/managers.

An understanding that the data will be used for payment purposes exerts a strong motivation to gather accurate data. Many countries are using coded data for hospital payment purposes, and some also use it for out-of-hospital care. In Australia the introduction of Activity Based Funding (ABF) has led to a large improvement in the attention and effort given to collecting the coded hospital discharge data. ABF has seen the elevation of clinical coding from the basement to the penthouse¹³. Senior managers become very interested in coding quality where the accuracy of the coding has a direct and measurable impact on the revenue of the hospital. However, this also brings with it a risk of perverse incentives to manipulate, mis-record, or otherwise tamper with the accuracy of the data. Often external auditing of the accuracy of the coding is used to mitigate these risks and ensure that payments are based on accurate data.

Using the data to measure the performance of the facility is one of the most beneficial purposes of a data collection. How well are we doing in meeting the needs of our community? How well are we doing compared to similar facilities? How does our facility compare to the national standards for our type of facility? These questions are fundamental to the proper management of each facility. However, as with payment systems, there is a risk of collecting inaccurate data if some facilities enhance or conceal performance indicators to make them appear better performers than others. Important use entails the need for greater care regarding data quality.

There is nothing that better focuses the attention of clinicians and managers on data validity, reliability and integrity than enabling them to compare their performance indicator data for this year with industry norms (best and worst performers) or with their performance last year. A by-product of this use of the data is to increase motivation and create a culture for data quality assurance and improvement.

Discussion

Moving from a vicious to a virtuous cycle – the factors that influence and support high quality data collections

Based on our observations of best practice, examples of the factors and actions needed to move from a vicious to a virtuous cycle can be divided into the following sections: the data, governance arrangements, collection process, infrastructure and the people - collectors and users.

Making **Data important** by establishing internal and external measures that are based on the quality and efficiency of care. These include Activity/funding and outcomes. A good example is the Best Practice Tariff in the Irish Hip Fracture Database where data quality is one of the standards required to earn the best practice tariff, along with clinical standards of care.¹⁴ Measurement of hospital acquired complications¹⁵ provides measurement of the quality of care based on the administrative data set. The clinical use of this data to reduce harm to patients by the clinicians focuses the clinical team on improving the quality of the clinical documentation making it easier for the data gatherers to produce an accurate record of each episode of care. Regular reviews by clinical teams with the data gatherers of clinical outcomes fosters a deeper understanding and importance of the data. An understanding of the efficiency of care, for example measurement of length of stay comparison by both clinicians and managers to understand the different uses of hospital resources by patient and clinicians fosters a deeper understanding of the importance of the data. Identifying and celebrating clinical champions or “evangelists” who have demonstrated the value of the coded data for clinical quality performance measurement and research fosters a powerful momentum for change towards improving the overall quality of the data.

The Irish National Office of Clinical Audit outlier process is an example of a use that merges the data checking process with clinical scrutiny. Where a suspected data error is identified, the first step is to check for a data artefact, and the second step is to check for a real clinical difference. The outcome is to create trust and buy-in to the audit with improvements in the data that can then be used elsewhere.

Make the **data Transparent** - standardise and create access to all internal stakeholders. The most rapid improvement in data quality occurred where the data was shared in a democratic way. Benchmarking portals on clinical outcomes, funding efficiency measures and costing, promote a focus on ensuring accurate data first, then attention is moved to improving actual performance. A good example of this is the Independent Health and Aged Care Pricing Authority in Australia’s¹⁶ national benchmarking portal. Initially this portal was only accessible internally. In 2022 costing data is now accessible by the public. Clinical outcome measures including low value care, patient safety, mortality, preventable health measures and compliance to national clinical standards are good examples of making the data transparent to measure variation in patient care.

Make the **data standard – count, code, and cost activity to a single standard**. Without a standard way to collect and classify patient data i.e., one patient journey is described in the same way by different service providers, it is difficult to aggregate or make any meaningful use of the data for funding benchmarking and public health planning purposes. In our research this foundational element was often overlooked. Examples of this were inconsistent ways of reporting dates, birth weights lacked a common unit of measure, additional diagnosis data was not indexed and sequenced accurately and codes were entered with additional spaces. This lack of compliance to a national standard makes it extremely difficult to aggregate data accurately. Compliance to national standards is a minimum requirement to count, code and cost in a meaningful way.

Best practice health systems have agreed minimum data set definitions. An example of this is the Australian Institute of Health and Welfare’s METOR¹⁷ definitions. However even health systems that are mature sometimes lack consistency in their counting definitions e.g., admission policy and rules governing changes that occur between care types in acute and sub-acute care.

We recommend the adoption of international classification standard and coding rules but be prepared to make local changes where there is a good reason. For example, the palliative care coding in ICD10AM 10th edition was identified as problematic in the National Audit of Hospital Mortality. Palliative care clinicians worked with the Hospital Pricing Office in Ireland to identify how to better reflect palliative care in HIPE18.

The Irish National Audit of Stroke is an example of changing the data set to produce better information and increase the buy-in of clinicians. The Irish National Audit of Stroke uses HIPE data and adds stroke audit-specific data items using HIPE portal additional screens. The Stroke audit did not use the HIPE discharge code because it could not distinguish between discharge to nursing home or to a nursing home that included a rehabilitation unit. HIPE added discharge mode variable in 2022 to identify reason for discharge to a facility¹⁹.

A **Governance** structure is ideally best administered by one body which is responsible for establishing the rules and compliance to the agreed coding, counting and costing conventions. An important function of this body is the continual maintenance and implementation of new editions of the standards. The establishment of expert groups to advise on national standards for medical records, application of coding standards and queries is an important input into the maintenance and improvement of the standards. A clinical advisory group performs an important function to ensure that clinicians have an input in the development of a national activity-based funding system through the provision of timely and quality clinical advice to inform pricing authority decision making in the development of the funding system. This independent body is also responsible for establishing the funding framework. Examples of these bodies are the Independent Hospital and Aged Care Pricing Authority in Australia, the Healthcare Pricing Office in Ireland and National Casemix Center of Excellence in Kingdom of Saudi Arabia.

An important aspect of governance is **establishing** and measuring national key performance indicators that are sensitive to variations in coding and counting variations ensures compliance to standards and builds a culture of continuous improvement. Examples of key performance indicators include submission deadlines, data quality scores on compliance to standards and results of external audits. Data quality performance indicators can form an important part of service level agreements between the payers and the providers.

Data Collection processes should be organised as a system, not on an individual health service/facility/clinic basis. In our work we found that in several cases coding was conducted in a sub-scale environment e.g., in rural or small hospitals. A sub-scale coding team may not have the size to have the appropriate skills or backup to manage vacancies or surges in work flow, particularly when the demand for skilled coding resources exceeds the capacity of the system to produce skilled workers. Coding should be conducted and optimised as part of a system, utilising abstracting/ reporting software that has the ability to automate simpler records/ natural language with regular reviews and refinements of maps from natural language to data categories. Workflows can be developed to take advantage of scale by coding by specialty/complexity mapped to competency of the coder. Technology and new data gathering processes will help improve the quality of data collection.

In some instances, we found that the **infrastructure** that was in place at the hospitals under review was not suitable for capturing the coded data in an efficient way. Entry of codes into spreadsheets, then re-entry into separate systems is an example of a process that can easily introduce errors into the data. IT systems need to have the capacity to accept the code structures and facilitate quality feedback loops to the data gatherers to ensure the data is gathered in accordance to the minimum data sets. National data repositories need to have “two-way” validation routines to ensure the data that is uploaded and consolidated has passed through a number of edit loops with feedback to those as close as possible to the source data.

The **people** that are involved in the data collection process need to be recognised for their output and valued. Quality data depends on a motivated appropriately skilled workforce. Key

characteristics of workforce planning include investing in building the capabilities of the existing workforce as well as newly trained staff. Characteristics of organisations that were approaching best practice included targeted competency-based training, continuous education, identifying champions, rewarding them and using them as change agents. Cross functional teams with clinicians and coders reviewing industry benchmarks and activity and safety data quality metrics to improve patient care were examples of best practice.

An area that was not understood well were common measures and methods to identify the appropriate work size and competency. Most work force decisions were based on rule-of-thumb measures and politics, rather than an evidence-based approach. The paradigm needs to be moved to adopt common methods based on service level standards and the appropriate competencies to code the complexity of patients at that facility. It may be useful to consider a regional hospital grouping model of coding production as an option rather than the on-site hospital model because coding services need to be of a minimum size in order to operate at optimum efficiency (where efficiency is dependent on both productivity and quality). Best practice research suggests the minimum size of a coding service (in addition to the Coding Service Manager) should be 7-8 (full time equivalent) coders. This size service allows for division of coding work between different levels of coder expertise, appropriate levels of coder support and supervision, on the job training and personal development, regular auditing and checking of coding quality and strategic and sustained promotion of the importance of coded data to clinicians.

Conclusions

Improving the quality of health service data so that it is sufficiently accurate, timely and consistent to make appropriate and reliable decisions is central to improving the availability and quality of health services. Yet even in those countries that are collecting and aggregating health services data, the quality of that data, is highly variable.

There are many factors that influence quality data collection including the timely use of trusted data by clinicians and other decision makers as close to the data source as possible. Clinicians are vital in the process of accurately recording what happens to a patient and to making the best use of the data.

All factors are not equal in their influence, in particular, in our view, standardising the data to enable aggregation is a fundamental foundational element, as is a motivated appropriately skilled workforce. An area that was not understood well were common measures and methods to identify the appropriate size and competency of the work force. Regular auditing, cross checking of the validity of the data from other sources and promoting the use of data with appropriate feedback loops are also important.

However, our hypothesis is that making the data important and transparent in funding and clinical care decision making by the clinicians and managers who use the data will ensure that resources and recognition are provided to the data gatherers, leading to a virtuous cycle of quality improvement.

References

¹ SCORE Global Report 2020 A visual Summary WHO <https://www.who.int/data/stories/score-global-report-2020---a-visual-summary>

² Guidance on a data quality framework for health and social care Health Information and Quality Authority <https://www.hiqa.ie/sites/default/files/2018-10/Guidance-for-a-data-quality-framework.pdf>

³ Exploratory Data Mining and Data Cleaning. Vol. 479. John Wiley & Sons. Dasu, Tamraparni, and Theodore Johnson. 2003.

-
- ⁴ Quality assessment of real-world data repositories across the data life cycle: A literature review J Am Med Inform Assoc. 2021 Jul; 28(7): 1591–1599. Published online 2021 Jan 26. doi: 10.1093/jamia/ocaa340 Siaw-Teng Liaw, Jason Guan Nan Guo, Sameera Ansari, Jitendra Jonnagaddala, Myron Anthony Godinho, Alder Jose Borelli, Jr, Simon de Lusignan, Daniel Capurro, Harshana Liyanage, Navreet Bhattal, Vicki Bennett, Jaclyn Chan, and Michael G Kahn
- ⁵ The Book of OHDSI Observational Health Data Sciences and Informatics the Book is a living document, community-maintained through open-source development tools, and evolves continuously. 2021-01-11 <https://ohdsi.github.io/TheBookOfOhdsi/>
- ⁶ capture, submit, process, analyse, and disseminate
- ⁷ Quality assessment of real-world data repositories across the data life cycle: A literature review Siaw-Teng Liaw, et.al J Am Med Inform Assoc. 2021 Jul; 28(7): 1591–1599. Published online 2021 Jan 26. doi: 10.1093/jamia/ocaa340
- ⁸ National Audit of Admitted Patient Information in Irish Acute Hospitals. Pavilion Health, Sept 2016 https://www.hpo.ie/latest_hipe_nprs_reports/Pavilion_Report_National_Audit_Admitted_Patient_Info_Sept_2016.pdf
- ⁹ PICQ® Performance Indicators of Coding Quality proprietary tool of Beamtree <https://beamtree.com.au/>
- ¹⁰ The Aisbett Method: An innovative way to provide statistically robust Clinical Coding Audits at a fraction of the cost of random sampling Chris Aisbett (deceased) Jack Aisbett, Cliona O’Donovan, Beth Reid, Mary E Black Paul O’Connor
- ¹¹ Best practice in the management of clinical coding services: insights from a project in the Republic of Ireland, Part 1. Reid B, Ridoutt L, O’Connor P, and Murphy D (2017) Health Information Management Journal, 46(2), 69-77. Best practice in the management of clinical coding services: insights from a project in the Republic of Ireland, Part 2. Health Information Management Journal, 46(3), 105-112 Reid B, Ridoutt L, O’Connor P, and Murphy D (2017)
- ¹² <https://humancapitalalliance.com.au/digital-health-unit/>
- ¹³ Health information management and clinical coding workforce issues Shephard J (2010). Health Information Management Journal 39 (3): 37–41.
- ¹⁴ Irish Hip Fracture Database National Report 2020 National Office for Clinical Audit (NOCA) https://s3-eu-west-1.amazonaws.com/noca-uploads/general/Irish_Hip_Fracture_Database_National_Report_2020_DESIGN_FINAL.pdf
- ¹⁵ Hospital-acquired complications (HACs) | Australian Commission on Safety and Quality in Health Care [Internet]. [cited 2019 Oct 21]. Available from: <https://www.safetyandquality.gov.au/our-work/indicators/hospital-acquired-complications>
- ¹⁶ National Benchmarking Portal (IHACPA), National bench portal <https://www.ihacpa.gov.au/resources/national-benchmarking-portal>
- ¹⁷ Meteor Metadata Online Registry Australian Institute of Health and Welfare (AIHW) <https://meteor.aihw.gov.au/content/344846>
- ¹⁸ Hospital In-Patient Enquiry (HIPE) Data Dictionary 2022 Version 14. Healthcare Pricing Office, Health Service Executive, Dublin, Ireland. HIPE palliative care flag at https://www.hpo.ie/hipe/hipe_data_dictionary/HIPE_Data_Dictionary_2022_V14.0.pdf
- ¹⁹ Hospital In-Patient Enquiry (HIPE) Data Dictionary 2022 Version 14. Healthcare Pricing Office, Health Service Executive, Dublin, Ireland HIPE discharge mode at https://www.hpo.ie/hipe/hipe_data_dictionary/HIPE_Data_Dictionary_2022_V14.0.pdf

THIS PAGE LEFT INTENTIONALLY BLANK