Information for intensive care evaluation: methods to assess and improve data quality and data processing

Arts, D.G.T.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 1

General introduction
Health care quality assessment

Due to budgetary constraints, insurance regulations and professional ambitions, quality assessment of health care is getting increased attention. Quality assessment can be defined as the critical appraisal of the measured results of a health care program in comparison with formulated objectives [1]. In quality assessment, outcome measures are very popular as quality indicators. Outcome measures can for example be patient satisfaction, quality of life, or mortality. A comparison of the measured quality of care between different time periods or between health care centers (benchmarking) can help to gain more insight into health care performance and to identify possibilities for improvement.

A specific field of health care where quality assessment is receiving increased attention is Intensive Care.

The concept of intensive care originated in the second half of the 20th century. Currently intensive care is defined as “a service for patients with potentially recoverable conditions who can benefit from more detailed observation and invasive treatment than can safely be provided in general wards or high dependency areas” [2]. The technologies to treat and monitor these patients and the requirement of a large number of specially trained personnel make intensive care relatively expensive. Consequently, even though intensive care does reduce mortality and morbidity for many patients, the effectiveness and efficiency remain topics of interest. Effectiveness can be defined as the measure of agreement between the objective and the result of a health care program for a defined population under average circumstances in general daily practice [3]. Efficiency of a health care program can be defined as the achievement of the objectives under an optimal use of resources [3].

To enable assessment of the quality of care, data is necessary, e.g. concerning the patient population and the delivered care. Several initiatives have therefore led to the setting up of medical registries enabling quality assessment in intensive care.

Data collection in medical registries

We define a medical registry as a systematic collection of a clearly defined set of health and demographic data for patients with specific health characteristics, held in a central database for a predefined purpose (based on Solomon et al.[4]). Specific patient characteristics (e.g. the presence of a disease or whether an intervention has taken place) determine which patients should be registered. Medical registries can serve different purposes, for instance as a tool to monitor and improve quality of care or as a resource for epidemiological research [5]. Examples of medical registries that have been set up for the assessment of the quality of intensive care can be found in the USA (IMPACT), the United
Kingdom (ICNARC) and in Austria (ASDI). In the Netherlands, the National Intensive Care Evaluation (NICE) registry was set up to enable the quality assessment of Dutch intensive care.

The NICE registry

By the initiative of Dutch intensive care physicians the NICE foundation was founded in 1996. The aim of the NICE foundation was to provide insight into the effectiveness and the efficiency of Dutch intensive care. To enable the assessment of these quality aspects of intensive care, data is needed. The NICE foundation therefore started a registry. The NICE registry contains data from all admissions in Dutch Intensive Care Units (ICUs) that participate in the project. Registered data includes patient characteristics, diagnoses, physiological data and laboratory data.

In some cases these data are collected automatically by extracting them from the electronic patient data management system (PDMS) [6] into a local registry database. Other ICUs collect the data manually by filling in case record forms (CRFs) that are manually entered into a local registry database. Each month the data from the local databases are transferred to a central registry database.

Quality of data in medical registries

To be useful, data in a medical registry must be of good quality. We do not want the measured quality indicator to be influenced by the quality of the registered data. In the context of a medical registry, data quality can be defined as “the totality of features and characteristics of a data set, that bear on its ability to satisfy the needs that result from the intended use of the data”[7].

Important characteristics that determine the quality of data are its accuracy (= conformity to the truth), its reliability (= measurements of individuals on different occasions, or of different observers, produce the same or similar results [8]) and its completeness (= extent to which all necessary data that could have been registered have actually been registered). However, in practice quite frequently wrong patients are registered or data items can be inaccurately recorded or not recorded at all [9-12]. To optimize the quality of medical registry data, participating centers can follow certain procedures designed to minimize inaccurate and incomplete data. For example automatic data checks can be performed on the registered data, or the data collectors can follow a special training.

Quality of data can be affected by differences in definitions. A study of Freer et al. [13] pointed out that concepts and definitions are used differently by various staff groups within one unit. For example people or institutions can vary in their definition and coding of the diagnosis ‘heart failure’. If organizations wish to
compare their performance by means of benchmarking, common data definitions are a prerequisite. Differences can be resolved by drawing sharp and clear definitions. In addition documentation of data can be standardized by using medical terminological systems.

**Standardized data by means of medical terminological systems**

The increased use of electronically stored medical data, for example in electronic patient records, has led to the need for structured and standardized registration of data. For this purpose several medical terminological systems have been developed. A terminological system is a system that interrelates concepts (e.g. medication or diagnoses) of a particular domain (e.g. intensive care) and provides their terms and codes [14]. The relations between the concepts within a terminological system can be hierarchical (e.g. Is-A) or non-hierarchical (e.g. has-location). In addition some terminological systems hold (formal) rules for the composition of new concepts by combining existing concepts. In literature, terms such as “terminology”, “thesaurus”, “vocabulary”, “nomenclature” and “classification” are often confused. De Keizer et al.[14] provide a description of a typology of terminological systems, including definitions and relationships between the above mentioned different types of terminological systems.

Several authors have specified required characteristics of terminological systems [15-17]. Among the most important characteristics of a terminological system are the completeness and the correctness of its content, i.e. the concepts, their terms, and the relations between the concepts. The content of a TS is of utmost importance for its acceptance. A physician needs to be able to be complete and sufficiently accurate in depicting the care process, and clinical researchers need to be able to be complete in selecting specific patient groups at any desired level of aggregation. To realize this all concepts, terms and relations belonging to the domain of the TS should be represented and should be correct. For example, we want sufficient terms attached to a concept, and we want the terms to be only the correct ones.

During the last decades many terminological systems have been developed with different domains and different structures ranging from strict hierarchies to complex semantic nets. Examples of these are the International Classification of Diseases (ICD) [18], the Systemized Nomenclature of Medicine (SNOMED) [19, 20], and the North American Nursing Diagnosis Association (NANDA) terminology [21]. Upon request of the Dutch NICE foundation a terminological system and corresponding software for the domain of intensive care (IC) have been developed. This terminological system is called Diagnoses for Intensive Care Evaluation, DICE [22].
The DICE terminological system
Quality assessment and the introduction of Patient Data Management Systems (PDMS) increased the need for structured and standardized registration of diagnostic information in Dutch intensive care. Until now there has hardly been any systematic and structured registration of reasons for ICU admission. This is largely attributed to the lack of an appropriate terminological system for describing these reasons for admission at the ICU. Therefore the DICE terminological system was developed. The DICE terminological system comprises reasons for admission to the Intensive Care Unit (ICU), and some of their characteristics, such as the anatomical localization, the dysfunction and the aetiology. DICE can be incorporated into a PDMS to 1) adequately describe the health status of a patient and 2) to enable aggregation of homogenous patient groups for the purpose of analysis and evaluation of intensive care. [22] Since the start of the DICE-project in 1997 DICE has developed to an extensive terminological system that currently contains 2373 concepts, of which 1456 are diagnoses that form reasons for admission to the ICU. To enable incorporation into a PDMS software has been developed. The DICE terminological system and accompanying software will soon be implemented at two Dutch intensive care units. Before DICE will be used in real practice its content needs to be checked for completeness and accuracy.

From data to quality indicators
Quality of health care is assessed by means of quality indicators, that are mostly related to the outcome of care. Outcome data can be subjective (e.g. patient satisfaction or quality of life) or objective (e.g. mortality, morbidity or length of stay).

Quality assessment projects within the field of intensive care often use the in-hospital mortality as quality indicator. However, death can result from many factors other than ineffective care. Mortality depends not only on the structure (e.g. equipment and staff) and the process of care (type and timing) but also on the input, e.g. severity of illness of the patient population [23]. Terminological systems such as DICE can help to account for this ‘case-mix’ by enabling aggregation of homogenous patient groups based on reason for ICU admission. In addition, several severity-of-illness scoring systems have been developed, that are frequently being used for case-mix correction in the intensive care population.

Scoring systems in intensive care
Scoring systems provide a quantitative measure of an individual patients’ severity of illness, based on characteristics that have been recognized as
important in increasing the risk of death. Scoring systems may either be anatomical or physiological [23]. Anatomical scoring systems assess the extent of injury, whereas physiological systems assess the impact of injury on function. Some examples of the latter are the Acute Physiology and Chronic Health Evaluation (APACHE) [24, 25] and the Simplified Acute Physiology Score (SAPS) [26] scoring systems, that were specifically developed for the general ICU population. Physiological scores may change during the ICU stay, as the physiological response varies. This aspect has been recognized in the development of the Sequential Organ Failure Assessment (SOFA) score [27]. The SOFA score provides a daily quantification of the amount of organ failure in intensive care patients.

Based on the severity of illness of an ICU population the expected mortality can be estimated, which can also be used in the correction for case-mix differences. Expected mortality is estimated by means of prognostic models.

Prognostic models

Prognostic models provide an estimate of the mortality in specific patient populations, based on characteristics of the patients. The observed mortality divided by the expected mortality in a population forms the Standardized Mortality Ratio (SMR). The SMR is used as a case-mix corrected indicator of quality of care. Which characteristics to include in a prognostic model and their importance (weight) is mostly determined through logistic regression. Other methods for development of prognostic models include Bayesian networks and decision trees. The APACHE and the SAPS prognostic models calculate the expected mortality by means of a logistic regression equation, which includes the severity-of-illness scores. Other well-known prognostic models that have been developed for the intensive care population are the Mortality Prediction Models (MPM) [28] and the Logistic Organ Dysfunction System (LODS) [29].

The development of a prognostic model requires a dataset that includes for a group of patients their individual characteristics and outcome. Whereas the development of a prognostic model is based on one patient population it will be applied to other populations, e.g. in a different time or in a different region or country. To assess the generalizability of a prognostic model, its performance needs to be validated on an external data set, different from the one that was used to develop the model. Such a measurement may cover several aspects of the performance of a prognostic model. They include the ability to distinguish between survivors and non-survivors (discrimination) and the extent to which the predicted mortality rates for groups of patients (calibration) or individuals (accuracy) are in concordance with the observed mortality rates.
Objectives and outline of this thesis

The general objective of the study described in this thesis is to evaluate and to assure the quality of the documentation of data in medical registries such as the NICE registry, and the quality of the processing of data for performance measurement. Regarding the quality of data documentation the following objectives can be distinguished:

- To identify causes of insufficient data quality and develop a framework of procedures for data quality assurance in medical registries (chapter 2).
- To assess the contribution of training in data definitions and data extraction guidelines to improve quality of data for use in intensive care scoring systems (chapter 3).
- To analyse the quality of data used to measure severity of illness in the Dutch National Intensive Care Evaluation (NICE) registry, after implementation of quality improving procedures (chapter 4).
- To evaluate the reliability and the accuracy of SOFA scores and to identify causes of errors in SOFA scoring (chapter 5).
- To obtain insight into methods for the evaluation of the quality of the content of terminological systems and to evaluate the content of the DICE terminology system (chapter 6).

Data processing is here refers to the transformation of the documented data into standardised mortality ratios as indicators for the quality assessment of care, by means of the prognostic models. Quality of data processing is hereby restricted to the assessment of the validity of the prognostic models that are used. Regarding the validation of prognostic models the following objective applies:

- To assess the effect of the size of the validation sample on the measured performance of prognostic models, and to assess the performance of these models for the ICU population in the NICE registry (chapter 7).
References
