Transforming Healthcare with the Internet of Things
J. Hofdijk et al. (Eds.)

doi:10.3233/978-1-61499-633-0-102

Automated Data Aggregation for Time-Series Analysis: Study Case on Anaesthesia Data Warehouse

Antoine LAMER[a, b, 1], Mathieu JEANNE[a,b], Grégoire FICHEUR[c], and Romaric MARCILLY[b]
[a] *Univ. Lille, CHU Lille, Pôle d'anesthésie-réanimation, F-59000 Lille, France*
[b] *Univ. Lille, Inserm, CHU Lille, CIC 1403 - Centre d'investigation clinique Innovations Technologiques, F-59000 Lille, France*
[c] *Univ. Lille, CHU Lille, EA 2694 - Santé publique : épidémiologie et qualité des soins, F-59000 Lille, France*

Abstract. Data stored in operational databases are not reusable directly. Aggregation modules are necessary to facilitate secondary use. They decrease volume of data while increasing the number of available information. In this paper, we present four automated engines of aggregation, integrated into an anaesthesia data warehouse. Four instances of clinical questions illustrate the use of those engines for various improvements of quality of care: duration of procedure, drug administration, assessment of hypotension and its related treatment.

Keywords. Data Collection, Vital Signs, Anesthesia.

Introduction

Operational databases daily collect high volumes of data in medical care centres and offer opportunities for reusing these data, e.g. for research purposes or in order to assess the quality of care [1]. For instance, analyses of anaesthesia time-series can be useful to identify the events that impact the patient outcome [2]. However, the structure of these databases is designed for clinical or administrative purposes, not for their secondary use [3]. Operational databases, in particular, collect high volumes of raw and heterogeneous data that cannot be reused directly in their initial format: for example, Anaesthesia Information Management Systems (AIMS) register data about the anaesthesia procedure mainly for medical follow up, and sometimes for legal purposes. Each anaesthetic procedure produces around two thousand measurements of vital signs (e.g. heart rate, arterial pressure) and one hundred events (e.g. surgical stages, administration of drugs). For clinical research purposes, it is important to assess specific events during predefined periods: for instance, the evolution of vital signs in relation to drug administration, the occurrence of some specific adverse events (e.g. hypotension, tachycardia) during anaesthesia or the cumulative dose of various drugs administered to the patient during the procedure. Despite the high volume of data collected by the AIMS, several quality

[1] Corresponding Author.

checks and various analyses must be carried out on these data before any clinical research can take place [4].

First, it is necessary to integrate data from heterogeneous sources into a common repository, such as a data warehouse: these well-known systems are used in order to manage high volume of data [4-5]. Then, after a cleaning and transformation process, data can be aggregated and synthesized in order to obtain relevant and meaningful information that can be used for clinical or statistical analyses. The aggregation process is based on business rules provided by expert clinicians.

The goal of this paper is to highlight opportunities related to data aggregation: turning raw data into substantial information for reuse. With this aim in mind, we describe four engines that have been developed to perform the automatic aggregation of anaesthesia time-series data and to carry out clinical or statistical analyses.

1. Methods

1.1. Available data

A data warehouse has been developed in the University Hospital of Lille to reuse data registered by the AIMS [6]. It stores data related to around 55 000 interventions performed each year (e.g. patient history, vital signs, various stages of the procedure, drug administration's) which then may be queried to fed aggregation engines.

1.2. Aggregation methods

The aggregation step consists in transforming data into meaningful information. *Data* are characterized by a high number of records (in rows) for a few numbers of attributes (in columns). On the contrary, *information* is characterized by a higher number of attributes (in columns) and a lower number of records. Therefore, the volume of data is decreased, while the volume of meaningful information is enhanced (Figure 1).

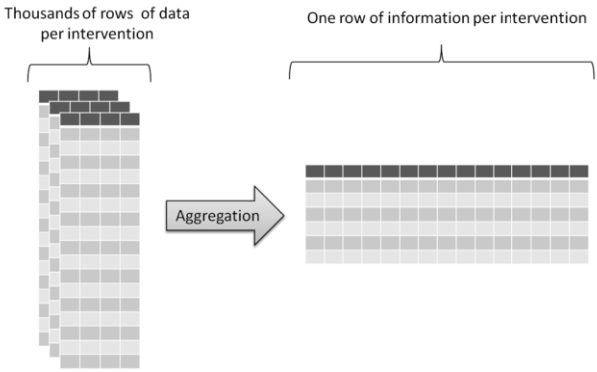

Figure 1. The aggregation process transforms raw data into aggregated data. Raw data are registered by source systems and present a high volume but a low information rate while aggregated data have a lower volume and more information.

We developed four aggregation engines to reuse anaesthesia time-series data (Figure 3). Each aggregation method was evaluated by experts by comparing results with source data. Engines can be used jointly according to the questions.

- The **"study periods"** engine is based on the detection and selection of predefined elements among the measurements, events, drugs (e.g. first non-null value, first administration of hypnotic drug) which correspond to the start or the end of specific periods (Figure 2, A1-A2) [7]. Example: anaesthesia or surgery stages, periods surrounding the administration of a given drug.
- The "**aggregated measures"** module employs aggregate functions (e.g. mean, median, minimum, maximum) to compute statistical indicators based on the measurements made during a certain study period previously defined (Figure 2, B1-B2). Example: mean of Heart Rate (HR) during anaesthesia, maximal value of mean arterial pressure within ten minutes after a vasopressor has been administered.
- The "**abnormal values of vital parameters"** module detects adverse events based on time-series of vital signs (Figure 2, C1-C2) [8]. These events are characterized by the total time elapsed outside a pre-defined range, the number of these episodes and the extreme (lowest or highest) value. Example: episodes of hypotension with Mean Arterial Pressure (MAP) lower than 60 mmHg.
- The "**drug administration"** module computes for each drug the total dose of the successive administrations within a study period (Figure 2, D1-D2). Example: total dose of propofol used during the induction period.

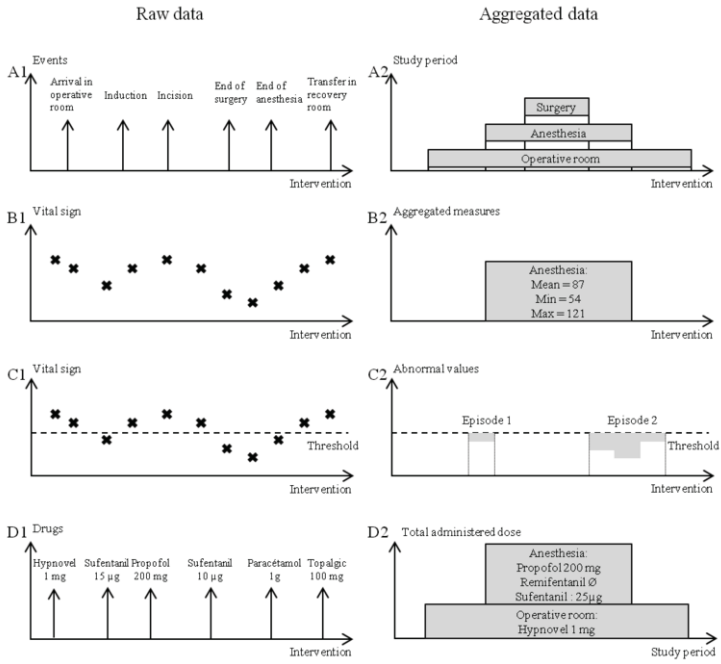

Figure 2. Aggregation engines. A) Study period ; B) Aggregated measures ; C) Adverse events ; D) Drug administration.

2. Results

In order to illustrate how the aggregated engines works, we have developed four study cases to know (i) what is the duration of the anaesthesia procedure, (ii) what are the variations of HR around administration of atropine, (iii) how can be characterized the drop of MAP after induction of anaesthesia with propofol and (iv) what is the total amount of ephedrine administered to manage blood pressure following the start of anaesthesia. From a clinical perspective, results for a given question can be regularly assessed over time to improve the quality of care A total of 276 775 anaesthetic procedures have been carried out at the University Hospital of Lille, France, between January 2010 and December 2014.

(i) Length of anaesthesia procedure - Engine "Study period": the length of anaesthesia has been measured for 261 996 anaesthetic procedures leading to a mean (standard deviation) of 161 (200) min (cf [7] for more details).

(ii) Evolution of heart rate - Engines "Aggregated measures" and "Study period": it has been calculated on 17118 interventions. HR evolves from a mean minimal value of 53 beats per minute (bpm) before administration of atropine, to a mean maximal value of 87 bpm after administration, and remains constant around 76 bpm after 15 minutes after the administration (figure 3).

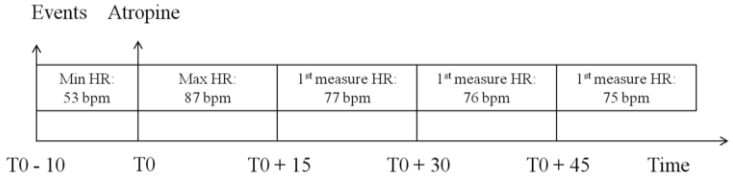

Figure 3. Aggregated measures of heart rate variations around the administration of atropine.

(iii) Abnormal values of Mean Arterial Pressure (MAP) - Engines "Study period" and "Abnormal values": it has been calculated on 81 014 interventions between induction and start of surgery. Table 1 presents, for various minimal thresholds of MAP, the number of interventions which dropped below the threshold, the mean duration of the first episode and the delay between the induction and the minimal value reached.

Table 1. Drop of Mean Arterial Pressure between induction of anaesthesia and start of surgery

Minimal Threshold (mmHg)	Nb of interventions (%)	Median duration of first episode (min)	Mean (SD) time between induction and minimal value (min)
< 50	11846 (14.62%)	3.05	16.91 (13.20)
< 55	10822 (13.36%)	2.98	17.93 (13.84)
< 60	14016 (17.3%)	3.67	18.17 (14.09)
< 65	14304 (17.66%)	4.87	17.94 (14.09)
< 70	11807 (14.57%)	4.95	17.79 (14.28)
< 75	8266 (10.2%)	4.97	17.67 (14.33)
-	9953 (12.29%)	-	-

(iv) Total dose of ephedrine - Engines "Study period", "Abnormal values" and "Drug administration": total dose of ephedrine, used as a treatment of hypotension

which may follow induction, has been calculated on 81 014 interventions. Table 2 presents, for each lowest thresholds of MAP, the number of interventions in which ephedrine has been administered between induction and start of surgery and the median (interquartile) of the dose administered.

Table 2. Administration of ephedrine between induction and incision

Threshold (mmHg)	Interventions with administration of ephedrine following induction (%)	Ephedrine (mg) (median [interquartile])
< 50	6566 (55.43%)	9 [9 ; 15]
< 55	4443 (41.06%)	9 [9 ; 12]
< 60	3747 (26.73%)	9 [6 ; 9]
< 65	1850 (12.93%)	9 [6 ; 9]
< 70	688 (5.83%)	9 [6 ; 9]
< 75	261 (3.16%)	9 [6 ; 9]
-	177 (1.78%)	9 [6 ; 12]
Total	17732 (21.89%)	9 [9 ;12]

3. Discussion

In this paper, we highlighted how aggregation engines can be used to permit reuse of data produced by the AIMS. The aggregation of raw data increases information while decreasing the volume of data. Even if developing these engines is time consuming and requires informatics and clinical expertise and validation, once engines are developed, they can be quickly and easily customized to carry out new studies. Moreover, the four engines presented in this paper can be adapted to various time-series (e.g. data recorded on patients in intensive care) and can be joined when it is necessary to combine various types of information to answer a specific clinical question. Routinely performed in a data warehouse, automated data aggregation makes aggregates available, even for records recently integrated, and directly operable for analysis and retrospective clinical studies. These studies can be repeated over time for quality of care assessment and improvement.

References

[1] M.G. Weiner, P.J. Embi. Toward Reuse of Clinical Data for Research and Quality Improvement: The End of the Beginning? Ann Intern Med. 2009 Sep 1;151(5):359–60.
[2] Sessler DI, Sigl JC, Kelley SD, Chamoun NG, Manberg PJ, Saager L, et al. Hospital stay and mortality are increased in patients having a “triple low” of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia. Anesthesiology. 2012 Jun;116(6):1195–203.
[3] K. Dentler, A. ten Teije, N. de Keizer, R. Cornet. Barriers to the reuse of routinely recorded clinical data: a field report. Stud Health Technol Inform. 2013;192:313–7.
[4] C.M. Nunez. Advanced techniques for anesthesia data analysis. Semin Anesth Perioper Med Pain. 2004 Jun;23(2):121–4.
[5] R. Kimball.The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing, and Deploying Data Warehouses. John Wiley & Sons (1998)
[6] A. Lamer, M. Jeanne, B. Vallet, G. Ditilyeu, F. Delaby, B. Tavernier, and R. Logier, “Development of an anesthesia data warehouse: Preliminary results,” IRBM, vol. 34, no. 6, pp. 376–378, décembre 2013.
[7] A. Lamer, J. De Jonckheere, R. Marcilly, B. Tavernier, B. Vallet, M. Jeanne et al. A substitution method to improve completeness of events documentation in anesthesia records. J Clin Monit Comput. 2015.
[8] A. Lamer, M. Jeanne, R. Marcilly et al. Methodology to automatically detect abnormal values of vital parameters in anesthesia time-series: proposal for an adaptable algorithm. Comput Methods Programs Biomed. 2015. In press.