International Journal of Pharmaceutics 505 (2016) 394–408

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

Knowledge management in secondary pharmaceutical manufacturing by mining of data historians—A proof-of-concept study

Natascia Meneghetti[a], Pierantonio Facco[a], Fabrizio Bezzo[a], Chrismono Himawan[b], Simeone Zomer[b], Massimiliano Barolo[a,*]

[a] CAPE-Lab—Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova PD, Italy
[b] GlaxoSmithKline R&D, Bld5, Park Road, Ware SG12 0DP, UK

ARTICLE INFO

Article history:
Received 11 September 2015
Received in revised form 25 February 2016
Accepted 20 March 2016
Available online 22 March 2016

Keywords:
Pharmaceutical manufacturing
Secondary manufacturing
Quality by design
Process analytical technology
PAT
QbD

ABSTRACT

In this proof-of-concept study, a methodology is proposed to systematically analyze large data historians of secondary pharmaceutical manufacturing systems using data mining techniques. The objective is to develop an approach enabling to automatically retrieve operation-relevant information that can assist the management in the periodic review of a manufactory system. The proposed methodology allows one to automatically perform three tasks: the identification of single batches within the entire data-sequence of the historical dataset, the identification of distinct operating phases within each batch, and the characterization of a batch with respect to an assigned multivariate set of operating characteristics. The approach is tested on a six-month dataset of a commercial-scale granulation/drying system, where several millions of data entries are recorded. The quality of results and the generality of the approach indicate that there is a strong potential for extending the method to even larger historical datasets and to different operations, thus making it an advanced PAT tool that can assist the implementation of continual improvement paradigms within a quality-by-design framework.

1. Introduction

In the last decade, the pharmaceutical industry has been faced with unprecedented business scenario changes. Many blockbuster drugs have been crossing the period of patent expiry and fewer blockbusters are on the horizon. The development of new products is shifting towards more complex therapeutic targets, and the patient base is narrower than that of preceding blockbusters (Kukura and Thien, 2011). Generic competition has become more and more aggressive (am Ende, 2011). Governments are taking radical measures to gain control over drug pricing (e.g. by changing the copayment plans; Sadat et al., 2014). Given this scenario, the pharmaceutical companies are striving to reduce costs to maintain competitiveness.

Primary pharmaceutical manufacturing is concerned with the production of active ingredients, whereas secondary pharmaceutical manufacturing focuses in the production of dosage forms (Bennett and Cole, 2003). Both primary and secondary manufacturing play a central role in cost allocation. However, while on the one hand the pharma industry is very effective in discovering new drugs, on the other hand its manufacturing efficiency is far behind the one of several other sectors. Poor performance in manufacturing costs the pharma industry US $90 billion per year, which is considered equivalent to the current development cost for 80–90 new drugs (The Economist, 2005; Danese and Constantinou, 2007). Based on the annual reports of 17 “big pharma” companies, it has been estimated that manufacturing costs amount to ~27% of the revenues, largely exceeding the R&D expenses that are at ~17% (am Ende, 2011). Therefore, even a fractional improvement in the quality of the manufacturing system can bring tremendous competitive advantages to a company.

Though product quality targets are very severe, pharmaceutical manufacturing processes still suffer for high variability. Continuous manufacturing is gaining more and more consideration, but most active pharmaceutical ingredients and drug products are still manufactured batchwise. Commercial manufacturing processes are often suboptimal, because they are conceived at the development stage and get frozen close to product registration, with little or no attempt to optimize them. Manufacturing cycle times are very variable, because out-of-specifications (“exceptions”) during

* Corresponding author.
E-mail address: max.barolo@unipd.it (M. Barolo).

http://dx.doi.org/10.1016/j.ijpharm.2016.03.035
0378-5173/

manufacturing need frequently be dealt with (Suresh and Basu, 2008). All of these factors contribute to significantly decrease productivity and increase product costs.

With the advent of fast, cheap and reliable on-line measurement devices, product manufacturing environments have now available large historical databases spanning several manufacturing years. However, while being data rich, the pharma industry is also known to be information poor (Politis and Rekkas, 2001). This is due to the fact that, due to data overload, the information embedded in data historians is hidden and therefore remains largely unexploited. Indeed, transforming data into knowledge is not a simple task. To clarify this issue, consider a typical secondary manufacturing system. The ingredients are processed by a series of batch operations, which eventually result in the final drug product. Each operation evolves through a series of phases, which may involve exchange of heat and/or mass with the surroundings and are often triggered by the operators. While a unit is processing the material, there may be short time windows where the unit is stalled (e.g. for re-setting, quick maintenance, and the like). At the conclusion of a batch, the equipment is possibly subject to maintenance and operation tests, then cleaned and set in a hold position for the next operation. Each piece of equipment is equipped with several sensors and hooked to a computer where sensor measurements (temperatures, flows, torques, compression forces, etc.) are recorded along with some settings (position of switches, controller set-points, etc.), for a total number of recordings on the order of a few tens at each time instant per piece of equipment. Typically, the recordings are made continuously (i.e., at the frequency of one set of recordings every few seconds) across an entire production campaign, which may last several months and may possibly include different products. In most cases, the data capture systems are meant to record data in a "passive" way only, i.e. without contextualizing the operations around them. Therefore, the recordings typically include also data segments that refer to temporary stalls of the equipment, where the time profiles of the recorded signals are totally unrelated to the evolution of the operation within the equipment; not even when the equipment is not processing material is the recording interrupted. The net result is that the amount of data records that gets archived for a given production campaign is overwhelming, easily reaching several millions of data entries. Additionally, the structure of the data capture systems may be out of step with respect to the implementation of newer and increasingly sophisticated data modeling and monitoring techniques, whose requirements were possibly not factored in at the time of the systems installation. A mechanical update of the systems to this end might even produce further disruption at significant cost for production.

Periodic review of the historical operational data by the company management is not easy, as the information is masked

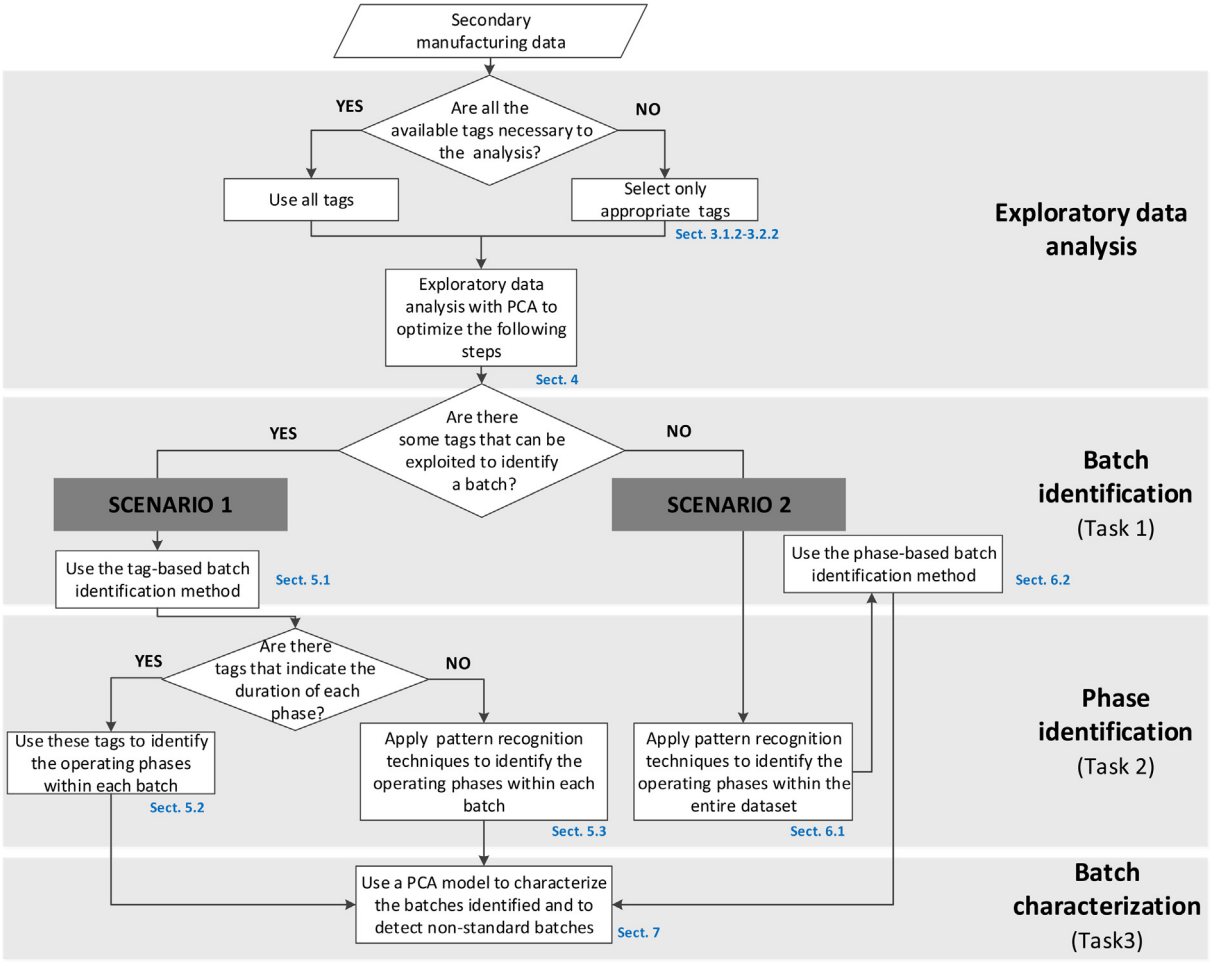

Fig. 1. Flowchart of the proposed approach to analyze secondary manufacturing data historians for batch systems. Each block includes a reference to the section where the block operations are discussed.