

# **Surgical Process Modelling: a review**

Florent Lalys, Pierre Jannin

INSERM, U1099, Rennes, F-35000, France

University of Rennes I, LTSI, Rennes, F-35000, France

*Corresponding author:*

Florent Lalys (florent.lalys@univ-rennes1.fr)

LTSI Faculté de médecine / Equipe U1099 Médecins  
Université Rennes I, 2 Av. du Pr Leon Bernard  
CS 34317, 35043 Rennes  
France

Tel.: +33 (2) 23 23 38 29

## ***Keywords***

Surgical workflow, procedural knowledge, surgical skill evaluation, computer assisted surgery

Word count:	9700
Line count:	830
Abstract word count	240
Tables:	4
Figures:	10
References:	95

## **Abstract**

Surgery is continuously subject to technological and medical innovations that are transforming daily surgical routines. In order to gain a better understanding and description of surgeries the field of Surgical Process Modelling (SPM) has recently emerged. The challenge is to support surgery through the quantitative analysis and understanding of Operating Room (OR) activities. Related surgical process models can then be introduced into a new generation of Computer-Assisted Surgery (CAS) systems and have a large impact in future surgical innovations, whether for planning, intra-operative or post-operative purposes. For instance they can improve surgical education and planning, situation awareness and management of complex multimodal information in the OR, surgical workflows, or increase surgical efficiency and quality of care. In this paper, we present a review of the literature dealing with SPM. This methodological review was obtained from a search using Google Scholar on the specific keywords: “surgical process analysis”, “surgical process model”, and “surgical workflow analysis”. This paper gives an overview of current approaches in the field that study the procedural aspects of surgery. We propose a classification of the domain that helps to summarise and describe the most important components of each paper we have reviewed, i.e. acquisition, modelling, analysis, application, and validation/evaluation. These 5 aspects are presented independently along with an exhaustive list of their possible instantiations taken from the studied publications. This review allows a greater understanding of the SPM field to be gained and introduces future related prospects.

# 1. Introduction

## 1.1 Context

In recent years, due to progress in Information Technology fields, computer assistance have been developed in healthcare systems, from hospital management through to medical imaging solutions. The Operating Room (OR), in particular, has undergone significant transformations evolving into a highly complex and technologically rich environment. Computer technologies are now essential and increasingly used throughout the intervention, from pre-operative planning to post-operative assessment. Computer-Assisted Surgery (CAS) (or Computer-assisted Intervention-CAI) systems have now a vital role to play in current surgical performance. For instance, during surgical planning, CAS and Image Guided Surgical systems provide access to multi-modal imaging technologies, relevant information about the patient and the possibility of simulating some parts of the surgical procedure. During surgery, they provide visualisation to pre- and intra-operative information about the patient with respect to the operative field, and also provide passive or active robotic support. New issues and technological challenges related to this complex OR and CAS systems have been discussed by Cleary et al. (2005), Rattner and Park (2003) or Xiao et al. (2008).

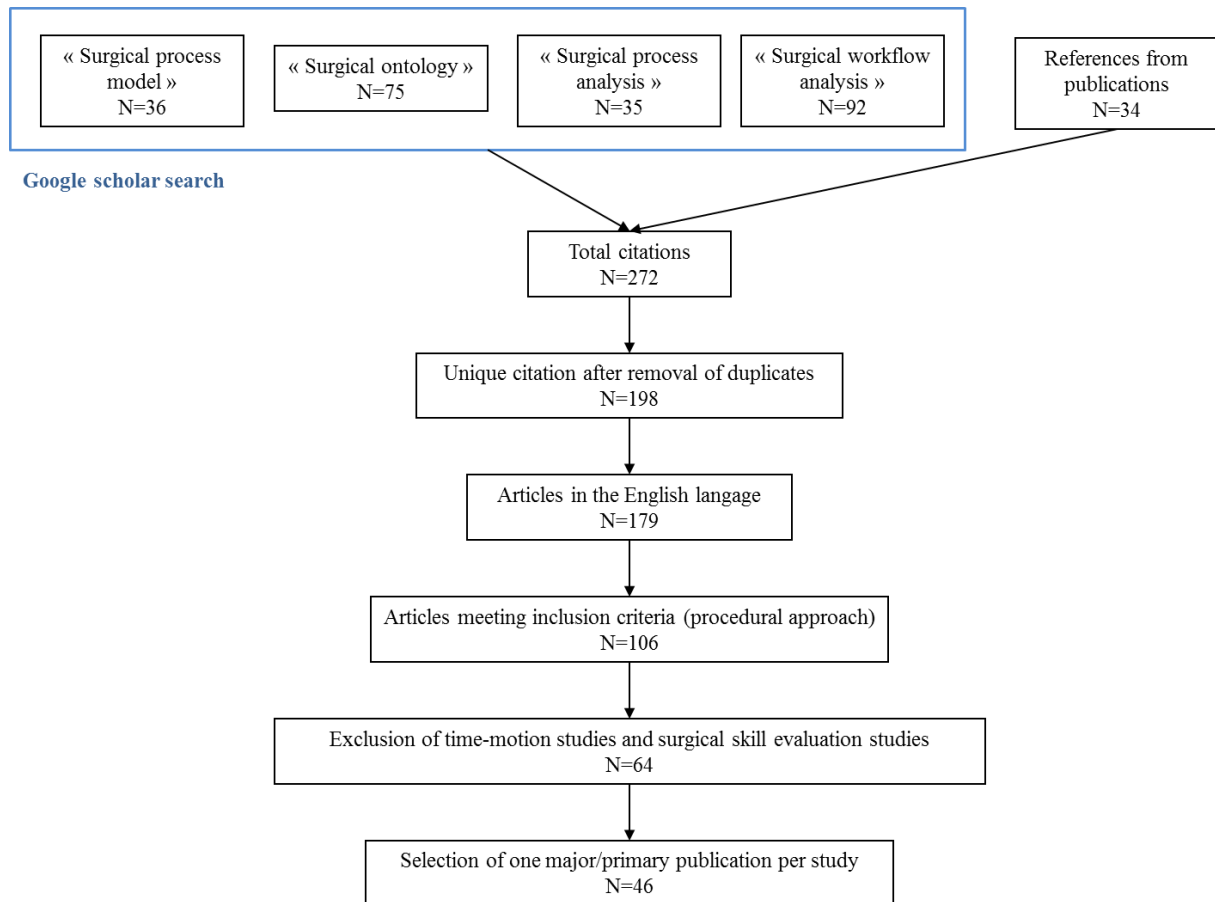
This first generation of CAS systems mainly focused on providing the surgeon with access to medical information of the patient before and during surgery, and active or semi-active robotic assistance. It was however outlined that such assistance would be different according to the surgical task, due to different needs and levels of importance. There was also an increasing need for new tools providing better resources management in the OR. For instance, surgeons need to be freed from technical problems through the automatic handling of software and hardware tools. These requirements illustrate the main motivation for surgical procedures models. The need for dedicated model-based systems for surgical procedures was first outlined for creating surgical simulation systems (Satava and Carrico, 1996; Taylor et al., 1999). The idea of describing the surgical procedure as a sequence of tasks was first introduced for analysis purposes in Minimally Invasive Surgeries (MIS) by MacKenzie et al. (2001), as well as for surgical planning and intra operative image management (Jannin et al., 2001) and for robotics systems (Munchenberg et al., 2000).

Following progress in model-based surgical intervention systems, understanding of the surgical scenario was proposed to improve the management of CAS systems. Jannin et al. (2003) defined the term *surgical model* as “generic or patient-specific surgical procedures that workflows aim to automate”. In their work, they stated that model-based systems of surgical interventions must address behavioural, anatomical and pathological aspects as well as integrate information about surgical instruments that can be used with *a priori* knowledge for the development of the OR of the future. The term *surgical workflow* has been defined as “the automation of a business process in the surgical management of patients, in whole or part, during which documents, information, images or tasks are passed from one participant to another for action, according to a set of procedural rules” (Jannin and Morandi, 2007). The term *Surgical Process* (SP) has been defined as “a set of one or more linked procedures or activities that collectively realise a surgical objective within the context of an organisational structure” (Neumuth et al., 2007). This term is generally used to describe the steps involved in a surgical procedure. A *Surgical Process Model* (SPM) has been defined as “a simplified pattern of an SP that reflects a predefined subset of interest of the SP in a formal or semi-formal representation” (Neumuth et al., 2007). It relates to the performance of an SP with support from a workflow management system. SPMs were first introduced for supporting surgical intervention using a model of surgical progress. Indeed, the precondition for computer-supported surgical intervention is the specification of the course model describing the operation to be performed. Typically, even if every surgery is different, a procedure of a same type shares common sequences of

states that can be identified. Such an assumption also makes it possible to detect abnormal activities within the procedure. Being able to identify information such as activities, steps or adverse events within a surgical intervention and having the possibility of relying on a surgical model is therefore a powerful tool in helping surgeons. The use of SPMs may prove effective in facilitating the surgical decision-making process as well as surgical teaching and assessment thereby having a direct impact on patient safety. It could help in anticipating patient positioning, optimising operating time, analysing technical requirements and improving the pre-operative human-computer interface. It could also help in evaluating surgeons or tools. In light of the growing interest in this field, and for the first time, we propose in this paper to undertake a methodological review of the literature focusing on the creation and the analysis of SPMs.

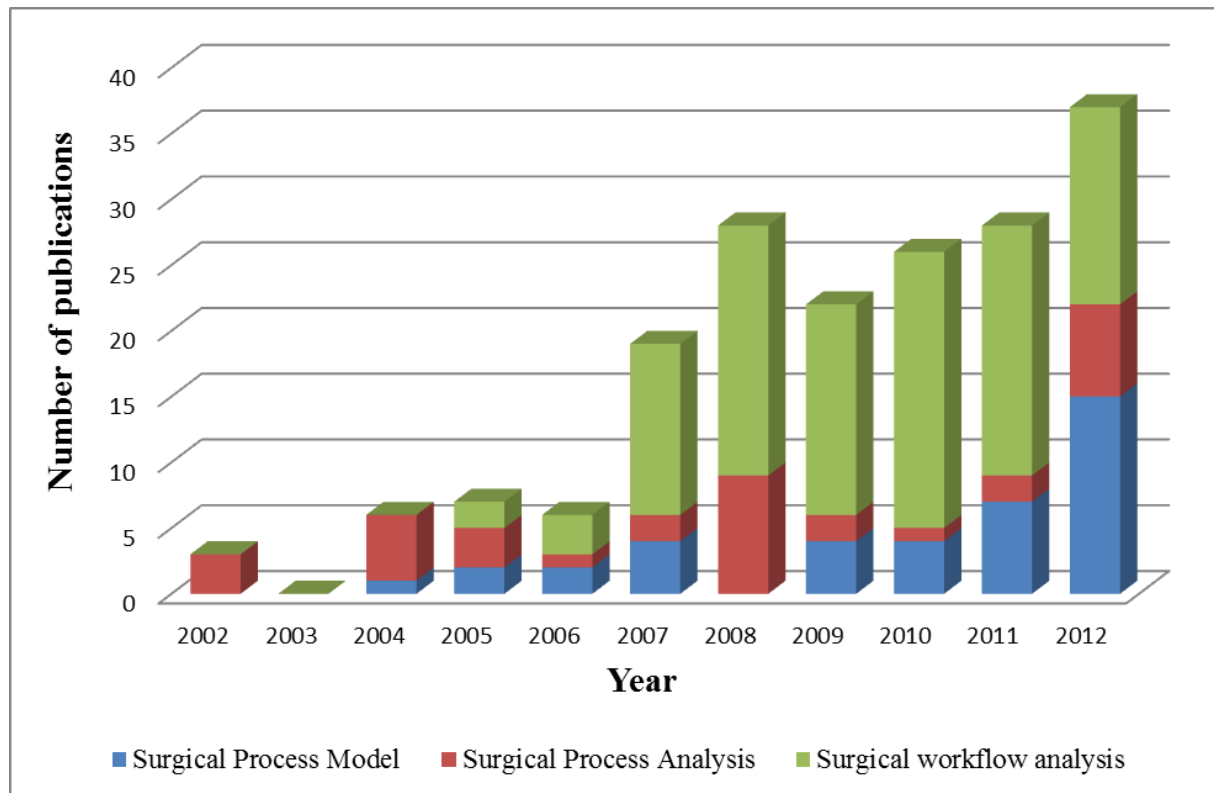
## 1.2 Search methodology

The review was carried out using Google Scholar to search on the specific keywords: “surgical process model”, “surgical process analysis”, and “surgical workflow analysis”. In addition to the Google Scholar results, we added another list of possible citations that were taken from the references of the first set of selected publications. We included articles published in peer-reviewed journals as well as full papers published in major international conference proceedings that dealt with the use of SPMs. International conference proceedings were included since the field is very recent, resulting in more conference publications than peer-reviewed journals. Only English language publications were selected. The research included was published between 2002 and end of 2012. In order to achieve an overview of the publications related to the creation and analysis of SPMs, we focused on publications which aimed to study the procedural dimension of surgery. The first inclusion criteria used during the selection process was therefore the fact that works have to take into account the sequential aspect of the surgical procedure, i.e. study the duration and sequencing of tasks performed during the surgery. Moreover, we were interested in pieces of work that focused at least one part of their analysis on the act of surgery, beginning when the surgeon performs the first task on the patient and ending when the surgeon makes the suture. It was defined as the second inclusion criteria. When a project has been published several times with no change in the dedicated elements of the diagram, the journal publication and if none, the most recent conference paper has been used. The entire selection process is shown on **Fig 1**. From an initial selection of N=272 publications, a total of N=46 publications were finally selected for full-text review.



**Fig 1** - Process used in the selection of publications for full-text review.

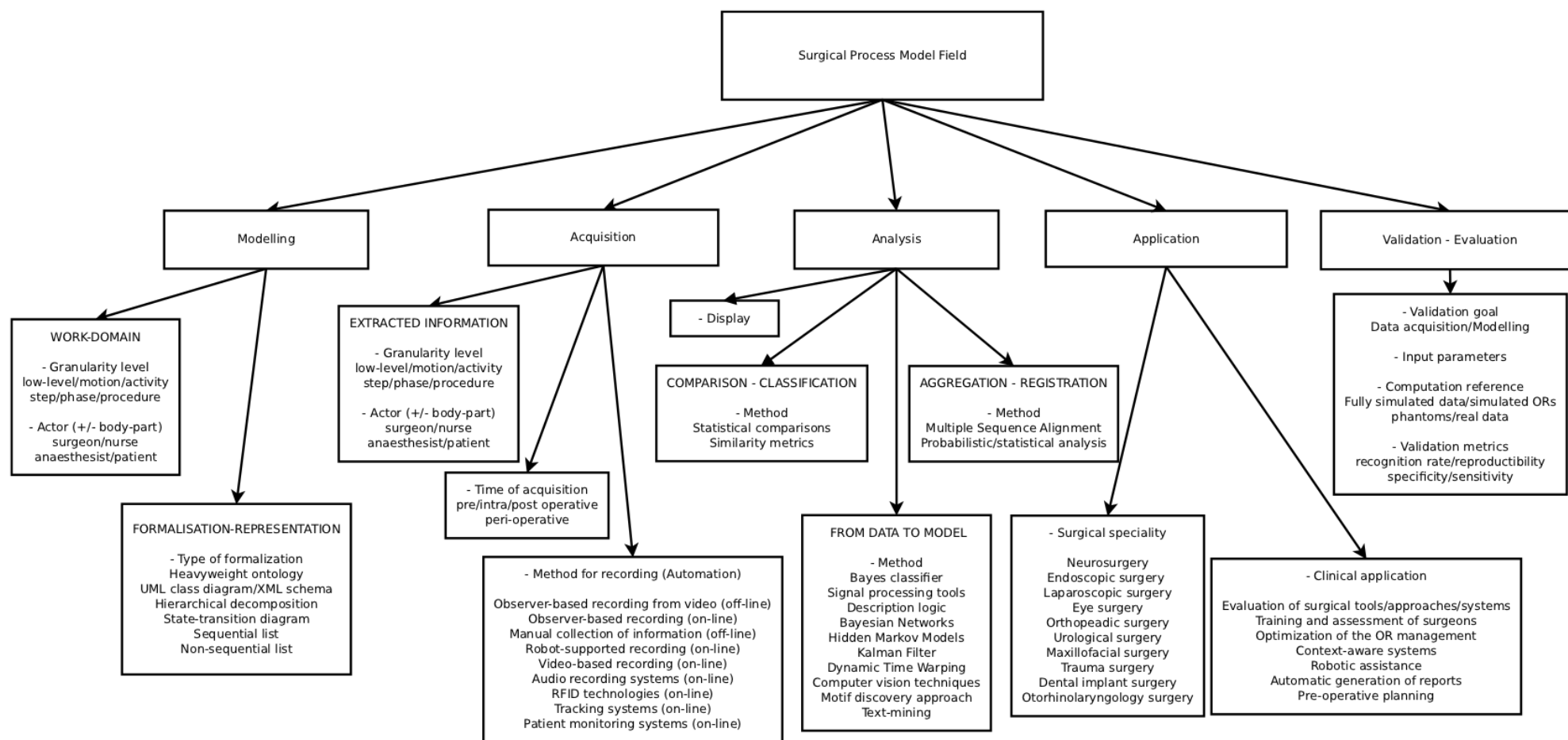
**Fig 2** shows the results of the Google scholar results before the selection process only. We can see that the SPM creation and analysis field is very recent. It has evolved in particular from 2007, evidence of the recent evolution of the field.



**Fig 2** - Evolution of the number of papers in the field from 2002 to December 2012.

## 2. SPM taxonomy

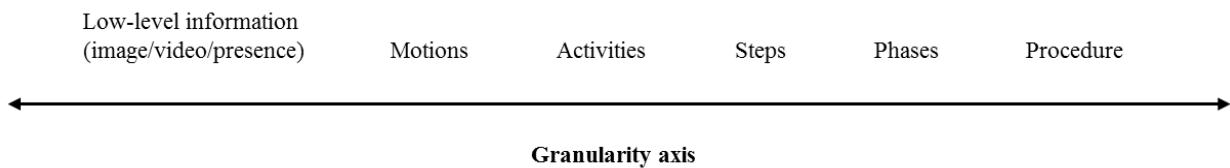
In order to clarify the review and the discussions, we propose a model for describing and classifying the methods using five components and their corresponding elements (**Fig 3**). Each of the five components addresses one major aspect of the SPM methodology, and every element that results can be instantiated with its set of possible values. The first component is the modelling, i.e. what is studied and is modelled, where the goal is to describe and formalize the work domain. The second component is the data acquisition performed by human observations or by sensor systems. The third is the analysis that tries to make the link between the acquired data and the information that we want to model. Another component specifies the different applications of the systems based on SPMs and finally the last component describes the different kind of validation and evaluation studies that were conducted for assessing these systems. The whole review is organised according to diagram of **Fig 3**. In the following subsections, each part of the diagram is explained in detail.



**Fig 3** - General overview of the field.

## 2.1 Granularity level

The whole SPM methodology, and especially the acquisition and modelling components, is organised around the concept of granularity level. A granularity level is defined as the level of abstraction at which the surgical procedure is described. New terms describing the different levels have been introduced and adapted to SPMs for improved standardisation of surgical descriptions. MacKenzie's group (Cao et al., 1996; Ibbotson et al., 1999; MacKenzie et al., 2001) proposed a model of the surgical procedure that consists of different levels of granularity: the procedure, the step, the substep, the task, the subtask, and the motion. Each (sub)task can for instance be broken down into various motions and forces primitives. They then used a hierarchical decomposition to structure the complex environment and the interaction between the surgical team and new technologies. Because of the marked differences in the terminology used in the papers studied, in this paper, we will use the following terminology for describing the different granularity levels of surgical procedures (**Fig 4**). The highest level is the procedure itself. The procedure is composed of a list of phases. A **phase** is similar to the notion of Lo et al.'s (2003) surgical episode, defined as the major types of events occurring during surgery. Each phase is composed of several steps. A **step** is considered to be a sequence of activities used to achieve a surgical objective. A step has been often called “task” in the literature. An **activity** is defined as a physical task. This level appears to be identical to a surgeme, previously defined as a well-defined surgical motion unit (Lin et al., 2006). Each activity is composed of a list of motions. The **motion** can be considered to be a surgical task involving only one hand trajectory but with no semantics. One assumption is that each granularity level describes the surgical procedure as a sequential list of events, except for the surgical procedure itself and for lower-levels where information may be continuous.



**Fig 4** - Different levels of granularity of a surgical procedure.

## 2.2 Modelling

This first component describes and explains the work-domain of the modelling, i.e. what is studied and what is modelled. Two elements are crucial for identifying the work-domain: 1) the granularity level at which the surgical procedure is studied and 2) the operator(s) involved in the surgical procedure on whom the study will focus. A third element can be added 3) formalisation. In many cases, a formalisation phase is required for representing the knowledge collected before the analysis process can take place. Knowledge acquisition is part of the underlying methodology of this component. It is the process of extracting, structuring and organising knowledge from human experts.

### Granularity level

Information that is studied (i.e. information that is modelled) is laid out on the granularity level axis defined in **Fig 4**. Investigations have concentrated on the activity level, but all granularity levels have been studied. At the highest level, the global procedure has been studied (Bhatia et al., 2007; Hu et al., 2006; Sandberg et al., 2005; Xiao et al., 2005), as well as the phases (Ahmadi et al., 2006; James et al., 2007; Katic et al., 2010; Klank et al., 2008; Lalys et al., 2012a; Lo et al., 2003; Nara et al., 2011; Padoy et al., 2007, 2008, 2010; Qi et al., 2006; Suzuki et al., 2012; Thiemjarus et al., 2012), the steps (Blum et al., 2008; Bouarfa et al., 2010; Fischer et al., 2005; Jannin et al., 2003, 2007; Ko et al., 2007; Lemke et al., 2004; Malarme et al., 2010), and the motions (Ahmadi et al., 2009; Lin et al., 2006; Nomm et al., 2008). Some studies integrated two or more of these granularity levels in their modelling (Burgert et al., 2006; Ibbotson et al., 1999; MacKenzie et al., 2001; Münchenberg et al., 2000; Xiao et al., 2005; Yoshimitsu et al. 2010). No low-level information has been considered here.

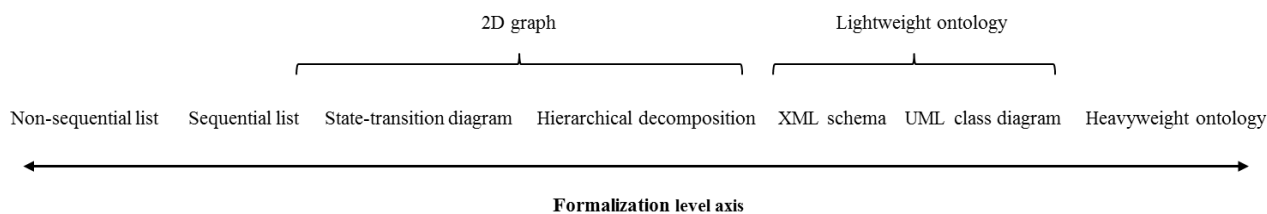


### Operator

The information that is studied involves one or more of the actors of the surgery: the operator studied may be the surgeon, the nurse, the anaesthetist, the patient or several of these operators.

### Formalisation

Formalisation is necessary for allowing automated handling and processing by computers. It is also necessary for bottom-up approaches to have a representation of the sequence of surgery through ontologies or a simple list of phases/steps/activities. At the highest level, we find the heavyweight ontologies, which have been used to represent the detailed context of a SPM study (Burgert et al., 2006; Fischer et al., 2005; Katic et al., 2010; Malarme et al., 2010; Speidel et al., 2008; Sudra et al., 2007). A heavyweight ontology is a lightweight ontology, i.e. an ontology based on a hierarchy of concepts and relations, enriched with axioms used to fix the semantic interpretation of concepts and relations. Then, in the category of lightweight ontologies, we find UML class diagrams and/or XML schemas (Jannin et al., 2003; Jannin et al., 2007; Meng et al., 2004; Neumuth et al., 2006b). Both approaches define entities and the relation between these entities. We then find all 2D graph representations, which have been used mostly, with hierarchical decompositions, state-transition diagrams and non-oriented graphs. Lastly, at the lower level, simple sequential (Agarwal et al., 2007; Ahmadi et al., 2006; Houliston et al., 2011; Hu et al., 2006; James et al., 2007; Klank et al., 2008; Nara et al., 2011; Padoy et al., 2007; Sandberg et al., 2005; Suzuki et al., 2012; Xiao et al., 2005) or non-sequential lists (Ahmadi et al., 2009; Lin et al., 2006; Nomm et al., 2008) were also used, suggesting a list of words for representing one or many levels of the surgery's granularity (**Fig 5**).



**Fig 5** - Different levels of formalisation of the surgery.

## **2.3 Data Acquisition**

The second component of the diagram, which is also the first step towards the creation of an SPM, is data acquisition, i.e. the collection of data on which the models are built. Four main elements may be distinguished in the acquisition process: 1) the level of granularity of the surgical information that is extracted, 2) the operator(s) on which the information is extracted, 3) the time when the acquisition is performed, and 4) the recording method. This section is divided according to these four elements.

### Granularity level

Like the Modelling component, the level of granularity of the surgical information that is extracted allows the acquisition to be characterised, as it determines in how much detail the SP is recorded. Studies have focused on the recording of the entire procedure (Sandberg et al., 2005), of the phases (Qi et al., 2006), of the steps (Burgert et al., 2006; Fischer et al., 2005; Lemke et al., 2004), of the activities (Forestier et al., 2012; Meng et al., 2004; Neumuth et al., 2006, 2009, 2012a, 2012b; Riffaud et al., 2011) and of the motions (Kragic and Hager, 2003). But efforts have been made in particular on extracting low-level information from the OR: videos (Bhatia et al., 2007; Blum et al., 2008; Haro et al., 2012; Klank et al., 2008; Lalys et al., 2012a, 2012b; Lo et al., 2003; Speidel et al., 2008), audio, position data (Houliston et al., 2011; Katic et al., 2010; Ko et al., 2007; Sudra et al., 2007), hand/tool/surgical staff trajectories (Ahmadi et al., 2009; Ibbotson et al., 1999; Lin et al., 2006; Miyawaki et al., 2005; Nara et al., 2011; Nomm et al., 2008; Yoshimitsu et al., 2010), information about the presence/absence of surgical tools (Ahmadi et al., 2006; Bouarfa et al., 2010; Padoy et al., 2007) or vital signs (Xiao et al., 2005). Several elements of this low-level information can also be combined (Agarwal et al., 2007; Hu et al., 2006; James et al., 2007; Malarme et al., 2010; Padoy et al., 2008, 2010;

Suzuki et al., 2012; Thiemjarus et al., 2012).

### Operator

Surgery always directly involves several operators. All staff members can have an impact on surgery and their roles and actions can be studied. The most important operator is of course the surgeon who is performing the surgery, but other operators can be involved: the nurse (Miyawaki et al., 2005; Yoshimitsu et al., 2010) for trajectory data extraction, the patient (Agarwal et al., 2007; Hu et al., 2006; Jannin et al., 2003, 2007; Münchenberg et al., 2000; Sandberg et al., 2005; Suzuki et al., 2012; Xiao et al., 2005) for images or vital signs extraction, or the anaesthetist (Houliston et al., 2011). Overall studies of the entire surgical staff have also been proposed (Agarwal et al., 2007; Bhatia et al., 2007; Fischer et al., 2005; Hu et al., 2006; Lemke et al., 2004; Nara et al., 2011; Qi et al., 2006; Sandberg et al., 2005; Suzuki et al., 2012), where the surgeon, the nurses and possibly the anaesthetist were included. For tracking systems, this notion can be specified by defining, in addition to the operator, parts of the human body involved such as the hand, eye, forehead, wrist, elbow and shoulder.

### Time of acquisition

The precise time of the data acquisition is also a vital piece of information for discriminating acquisition techniques. In most of the studies, data are extracted from intra-operative recordings. In some studies, this was done post-operatively (*retrospective*), for instance when recordings are performed by an observer from video, or in the case of certain tracking systems. In the case of the manual collection of information, this is done pre-operatively (*prospective*). Additionally, the term peri-operative generally refers to the three phases of surgery. Some acquisitions include all of these three phases to obtain information about the entire patient hospitalisation process (Agarwal et al., 2007; Sandberg et al., 2005).

### Recording methods

Two main approaches have been proposed: observer-based and sensor-based approaches (**Tab 1**). Observer-based approaches are performed by a human observer. For off-line recording, the observer uses one or multiple videos from the OR to record retrospectively the surgical procedure (Ahmadi et al., 2006, 2009; Bouarfa et al., 2010; Fischer et al., 2005; Ibbotson et al., 1999; Lemke et al., 2004; MacKenzie et al., 2001; Malarme et al., 2010; Padoy et al., 2007). For on-line recording, the observer is directly in the OR during the intervention (Forestier et al., 2012; Neumuth et al., 2006a, 2006b, 2009, 2012b; Rifaud et al., 2011). Lemke et al. (2004) first highlighted the importance of studying OR processes using on-line observer-based approaches to study both ergonomic and health economic aspects.

Sensor-based approaches have been developed for automating the data acquisition process and/or for finer granularity descriptions. The principle is to extract information from the OR using one or multiple sensors in an automatic way, and to recognise activities or events based on these signals. Sensors can be of different types, ranging from electrical to optical systems. In the beginning, studies used sensors based on Radio Frequency IDentification (RFID) technologies, directly positioned on instruments or on the surgical staff during the intervention, to detect the presence/absence of the tools or actors (Agarwal et al., 2007; Houliston et al., 2009; Neumuth and Weissner, 2012c). Then, efforts were made to use robot-supported recording (Ko et al., 2007; Kragic and Hager, 2003; Lin et al., 2006; Münchenberg et al., 2000), including surgeon's movements and the use of instruments. Robots have been used as a tool for automatic low-level information recording. Tracking systems (Ahmadi et al., 2009; James et al., 2007; Katic et al., 2010; Miyawaki et al., 2005; Nara et al., 2011; Nomm et al., 2008; Sudra et al., 2008; Thiemjarus et al., 2012; Yoshimitsu et al., 2010) have also been used in various studies; mainly through eye-gaze tracking systems positioned on surgeons or through staff-member tracking devices. Other types of methods have also been tested for recording information: patient monitoring systems (Agarwal et al., 2007; Hu et al., 2006; Sandberg et al., 2005; Xiao et al., 2005), and audio recording systems (Agarwal et al., 2007; Suzuki et al., 2012). Finally, the use of on-line video-based recording, sometimes combined with other data acquisition techniques, has especially received increased attention recently (Bhatia et al., 2007; Blum et al., 2008; Hu et al., 2006; James et al., 2007; Klank et al., 2008; Lalys et al., 2012a, 2012b; Lo et al., 2003; Padoy et al., 2008, 2010; Speidel et al., 2008; Suzuki et al., 2012), with either wide-angle video camera recording the entire OR or surgical video camera such as endoscopic or surgical microscope videos.

Observer-based approaches			Sensor-based approaches					
Observer-based recording from video (off-line)	Observer-based recording (on-line)	Manual collection of information (off-line)	Robot-supported recording (on-line)	Video-based recording (on-line)	Patient monitoring systems (on-line)	RFID technologies (on-line)	Tracking systems (on-line)	Audio recording systems (on-line)

**Tab 1-** List of possible data acquisition methods.

## 2.4 Analysis

Analysis methods can be divided into three types: methods that go from data to final model, methods that aggregate or fuse information and methods that classify or compare data to extract a specific parameter. The three approaches are presented in the following subsections. Additionally, methods for displaying the analysis results have been studied to obtain a visual representation after the analysis process.

### From data to model

The challenge here is to use the data collected during the acquisition process to create an individual model (i.e. iSPM) and to make the link between the acquisition process and the modelling. The type of approach used can be determined by comparing the level of granularity of the acquisition information and of the modelling. Top-down approaches are defined as analyses that start from a global overview of the intervention using patient-specific information and a description of high-level tasks (such as phases or steps) to fine-coarse details (such as activities or motions). Conversely, bottom-up approaches use as their input low-level information from sensor devices and try to extract high-level semantic information. The methodology employed for either bridging the semantic gap in the case of bottom-up approaches or to generalise and formalise individual recordings in the case of top-down approaches, is based on statistical or data-mining concepts. The level of automation during the creation of the model has to be defined. The issue is to determine whether or not the model needs a training step. This step is needed for assigning classes to the training set. In such cases, the creation of the model is not fully automatic and may be entirely manual or a mix between human intervention and automatic computation.

As part of supervised approaches, simple Bayes classifier with Linear Discriminant Analysis (Lin et al., 2003) and neural networks (Houlston et al., 2011; Nomm et al., 2008) have been tested in the case of activity/step/phase recognition. Signal processing tools have been used for analysing patient vital signs (Hu et al., 2006; Xiao et al., 2005) or audio recordings (Suzuki et al., 2012). In the case of *top-down analysis*, description logic has been tested (Burgert et al., 2006; Fischer et al., 2005; Katic et al., 2010; Lemke et al., 2004; Sudra et al., 2007), as well as model instantiation (Jannin et al., 2003), decision tree (Jannin et al., 2007), inference engine (Malarme et al., 2005) or workflow engine (Qi et al., 2008). In the case of *bottom-up analysis*, graphical probabilistic models have often been used to describe dependencies between observations. Bayesian Networks (BN) have recently proven to be of great interest for such applications, with an extension in the temporal domain using Dynamic BNs (DBN). Temporal modelling allows the duration of each step and of the entire process during its execution to be evaluated. Many time-series models, such as Hidden Markov Models (HMM) (Rabiner, 1989) or Kalman filter models, are particular examples of DBNs. Indeed, HMM, which are statistical models used for modelling non-stationary vector times-series, have been widely used in SPM analysis (Bhatia et al., 2007; Blum et al., 2008; Bouarfa et al., 2010). The Dynamic Time Warping (DTW) algorithm has also been often tested with success because of its ability precisely wrap time-series (Ahmadi et al., 2006; Padoy et al., 2008). Computer vision techniques, that allow a progression from a low-level description of images and videos to high-level semantic meaning, have also been used for extracting high-level information before using supervised approaches such as neural networks (James et al., 2007), Support Vector Machines (Support Vector Machines) (Klank et al., 2008), Bayesian networks (Lo et al., 2003), HMMs/DTW (Lalys et al., 2012a, 2012b; Padoy et al., 2008, 2010) or Linear Dynamical System, spatio-temporal features and multiple kernel learning (Haro et al., 2012). Computer vision techniques have also been mixed with description logic (Speidel et al., 2007). SVMs have been employed before the use of time-series analysis (Bhatia et al., 2007). Statistical analysis (Agarwal et al., 2007), sequential analysis (Ko et al., 2007; Kragic and Hager, 2003; Münchenberg et al., 2000), trajectories data mining (Nara

et al., 2011), times automata (Yoshimitsu et al., 2010) or model checking (Miyawaki et al., 2005) have also been used. Dealing with heterogeneous sources, a multi-objective Bayesian framework has finally been implemented for feature selection, and supervised classifiers were then launched (Thiemjarus et al., 2012).

As part of unsupervised approaches, no extensive work has been undertaken. Only a motif discovery approach has been used (Ahmadi et al., 2009) that does not need any *a priori* model.

Finally, an SPM whose data acquisition and modelling stay at the same level of granularity is also possible. In such cases, the goal of the analysis is not to create a real model, but to perform either aggregation/fusion or comparison/classification.

#### Aggregation-Fusion

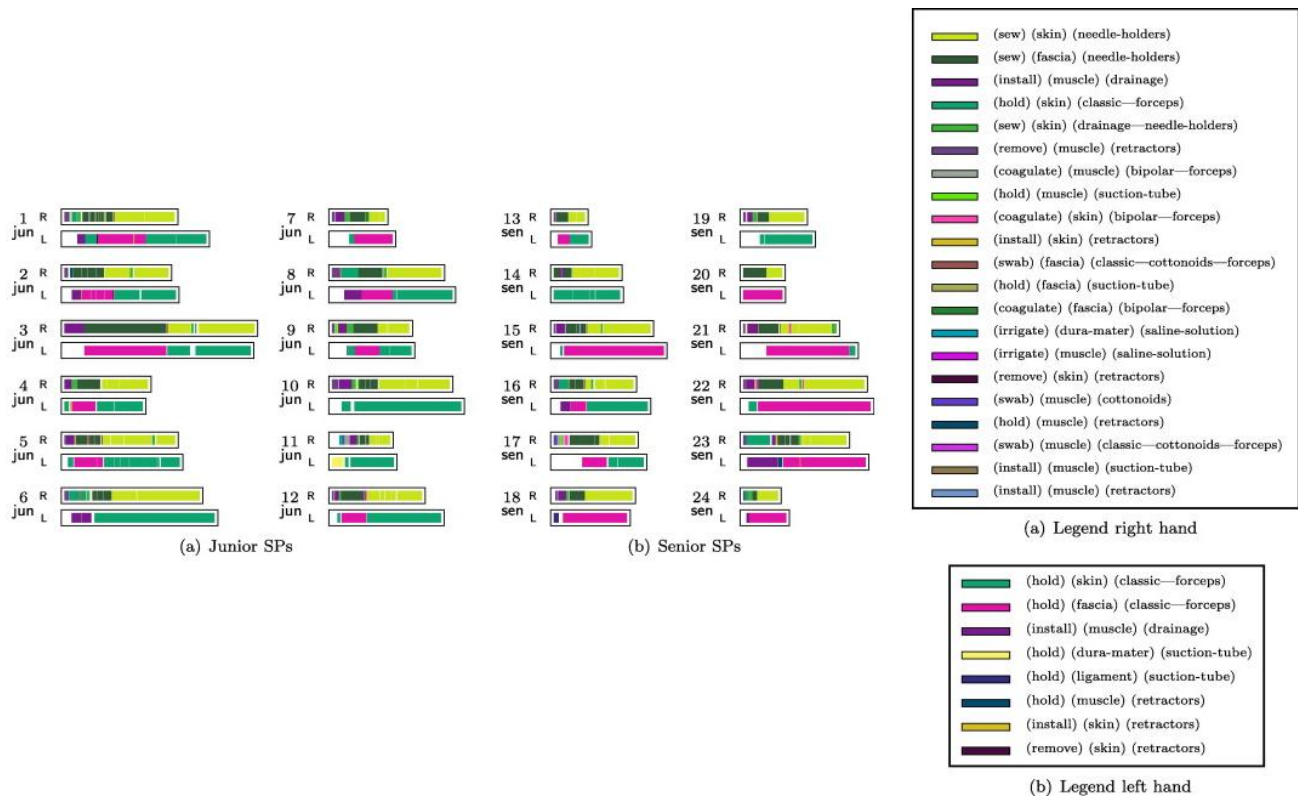
The goal here is to create a global (i.e. generic) model (gSPM) of a specific procedure representing a population of surgical procedures by merging a set of SPMs. One possibility is to merge similar sequences as well as filter infrequent ones to create average SPs in order to obtain a global overview of the surgery. Another is to create gSPMs that represent all possible transitions within SPs. A synchronisation stage may be necessary for both approaches in order to be able to merge all SPs. Generally, probabilistic or statistical analyses have been used for the fusion (MacKenzie et al., 2001; Neumuth et al., 2006b), but Multiple Sequence Alignment has also been tested (Meng et al., 2004) within text-mining approaches for automatically analyse post-operative procedure reports as well as patient files.

#### Comparison-Classification

The principle is to use an SPM methodology to highlight a specific parameter (i.e. meta-information) that explains differences between populations of patients, surgeons, or systems. Simple statistical comparisons (such as average, number of occurrence or standard deviation) have been used (Ibbotson et al., 1999; Riffaud et al., 2010; Sandberg et al., 2005) to compare populations. Similarity metrics have also been proposed by Neumuth et al. (2012a) to compare different SPs. DTW along with the K-Nearest Neighbour (KNN) algorithm have been tested within unsupervised approaches (Forestier et al., 2012).

#### Display

Once data is acquired and the model is designed, it is generally useful to have a visual representation of the data to explore them qualitatively and to illustrate the results. However, complex data structures sometimes prevent straightforward visualisation. High-level task recordings of SPMs can be displayed according to two types of visualisations: temporal and sequential aspects (Neumuth et al., 2006a). Temporal display focuses on the duration of each action, whereas sequential display focuses on the relation between work steps. Moreover, in the sequential display, one possibility is to create an exhaustive tree of each possible sequence of work steps. Sensor-based recordings are easier to visualise. As it is represented by time-series data, an index-plot can be used (e.g. in Forestier et al., 2012). The idea of an index-plot is to display the sequence by representing an activity as a rectangle of a specific colour for each value, and a width proportional to its duration. An information sequence can be easily visualised and a quick visual comparison can be performed (**Fig 6**).



**Fig 6** – Index-plot used in Forestier et al. (2012) representing the activities of the right (R) and left (L) hand for a population of 24 lumbar disk herniation surgeries performed by junior (a) and senior (b) surgeons.

## 2.5 Clinical applications

The analysis and modelling of surgical procedures cover multiple surgical specialities, issues, and challenges. Five major applications in particular have been the focus of increased attention: 1) evaluation of surgical tools/systems/approaches, 2) training and assessment of surgeons 3) optimisation of OR management 4) context-aware systems, and 5) robotic assistance. We first present the surgical specialities that have been covered by these systems, and then the five main applications are detailed. A final subsection presents other potential applications.

### Surgical speciality

SPMs have been applied to many surgical specialities, but Minimally Invasive Surgery (MIS), including endoscopic and laparoscopic procedures and neurosurgical procedures have been preferred. Within laparoscopic and endoscopic procedures, Cholecystectomy and Functional Endoscopic Sinus Surgery (FESS) surgeries have been widely studied. Works can also be found in eye surgery (Lalys et al., 2012a, 2012b; Neumuth et al., 2006b, 2012a, 2012b), maxillofacial surgery (Münchenberg et al., 2000), trauma surgery (Agarwal et al., 2007; Bhatia et al., 2007; Xiao et al., 2005), dental implant surgery (Katic et al., 2010), urological surgery (Meng et al., 2004), and otorhinolaryngology (ORL) surgery (Neumuth et al., 2006b). In general, systems have been specific to a surgical speciality or even a particular surgical intervention, but a few papers have described more generic surgical systems.

### Applications

*Evaluation of tools/surgical approach/systems:* The evaluation of surgical tools or systems was the first application targeted by research laboratories, at the request of surgeons (Ibbotson et al., 1999; Lemke et al., 2004; MacKenzie et al., 2001; Meng et al., 2004; Neumuth et al., 2006b, 2009; 2012b; Weinger et al., 1994). The analysis methods used in such cases were the comparison and classification methods.

*Training and assessment of surgeons:* All junior surgeons currently train with the teaching help of senior surgeons. This is a very time-consuming, interactive and subjective task. Moreover, there is growing pressure on surgeons to demonstrate their competences. The need for new automatic training systems using tools for the evaluation of surgeons

has motivated extensive research into the objective assessment of surgical skills (Hager et al., 2006, Rosen et al., 2001). It would allow surgeons to benefit from constructive feedback, and to learn from their mistakes. Surgical expertise has been widely studied in the literature. It is usual to distinguish technical from non-technical skills (Yule et al., 2006). Technical skills include motor skills as well as procedural and conceptual knowledge (Patel et al., 2001). Non-technical skills include cognitive skills and interpersonal skills (Yule et al., 2006). Surgical process modelling is a methodology which allows some aspects of motor skills (timing or trial-error loops, for instance) and some aspects of procedural knowledge to be assessed. The ability to recognise simple movements, activities, steps, or phases precisely is a very powerful tool in automating surgical assessment. Similar methods can also be used for training and assessing other members of the surgical team. Surgical training may also benefit from SPM methodology since it allows access to a formal description of the entire procedure, or a possible surgical scenario inside a population of cases (as represented by gSPM). For a complete discussion on the motivations of objective skill evaluation, one can refer to Reiley et al. (2011).

*Optimisation of OR management:* The need for perioperative surgical workflow optimisation has recently emerged (Dexter et al., 2004; Hu et al., 2006), especially regarding the specifications of the OR of the future (Cleary et al., 2005). With the increased number of CAS systems and new technologies, being able to manage and coordinate all these systems correctly is becoming vital. The optimisation of the use of physical and human resources required in an OR suite can reduce efforts and therefore improve patient outcomes, reduce hospital's costs and increase efficiency. Moreover, being able to identify different phases within the OR could be useful to know how to assign staff, prepare patients or prioritise OR clean-ups. Additionally, there are some adverse events that need to be taken into account. These may be long surgical interventions or emergencies that require the use of the OR without prior planning.

*Context-aware systems:* Many CAS systems, such as Augmented Reality (AR) systems or new imaging protocols, have been developed recently and integrated in the OR. Some limitations have been outlined. They are mainly used for a short period of time only, and the visualisation of additional information strongly depends on the current state of the intervention. Moreover, surgeons have to deal with adverse events during operations, arising from the patient him/herself but also from the management of the operation. The idea is to be aware of the current surgical situation in order to adapt assistance accordingly (e.g. in Sudra et al., 2007). Additionally, difficulties can be detected and risk situations better handled. For instance, variations of live signals can be used to warn surgical staff in the detection of anomalies.

*Robotic assistance:* Many pieces of research have demonstrated the importance of robots in assisting surgery, and particularly using SPMs (Ko et al., 2007; Kragic and Hager, 2003; Münchenberg et al., 2000; Miyawaki et al., 2005). Surgical robots play a vital role in improving accuracy in surgical procedures. Two families of robots have been introduced for intra-operative assistance: semi-active and active robots. Semi-active robots make the link between surgeon and patient. Surgeons perform their tasks outside the OR using the robot which reproduces the surgeon's hand movement on the patient. These types of robots are used for specific tasks only such as biopsies or endoscopies for MIS. Active robots are used directly in the OR to replace the surgeon for certain tasks. Both types of robots could benefit from SPMs in supporting these tasks using pre-defined models. The use of robotic assistance also aims to compensate for the lack of human resources in many hospitals, and in particular the lack of nurses (Miyawaki et al., 2005; Yoshimitsu et al., 2010). The new generation of robots that are currently being tested are able to pinpoint the progress of the intervention by automatically acquiring data from the surgical environment and creating SPMs.

Two other applications that have often been implicit in multiple publications are the automatic generation of post-operative reports and the help in pre-operative planning.

Post-operative reports are paper or electronic files that are generated post-operatively by the surgeon for documenting surgical procedures. Procedures are described as a succession of actions and steps that are manually included in a "log-file" for further filing. This step of the procedure is very tedious and time-consuming. The idea behind automating this process is to automatically extract as much information as possible from the surgery with the help of multiple sensors and to create pre-filled reports (Coles and Slavin, 1976; Lalys et al., 2012a). All studies that retrieve information from the OR, regardless of their level of granularity have potentially the possibility of automatically creating pre-filled reports.

For helping pre-operative planning, the goal is to better anticipate adverse events and possible problems during surgery by using formalised knowledge acquired by previous interventions and also by having an idea of all the possibilities offered by SPs. Aggregation and fusion techniques may be helpful in such cases for creating gSPMs.

## 2.6 Validation - Evaluation

We distinguish validation, defined as studying whether the system or method is actually doing what it is intended to do, from evaluation, defined as the study of the added value of a system or a method. Each aspect of the SPM methodology is subject to validation. The design of a validation study includes 1) the specification of a validation goal, 2) the definition of input parameters, 3) the computation or estimation of a reference (validation data sets) against which the results of the method to be validated will be compared, 4) the definition of validation metrics that will quantify the comparison, and 5) the operator using the system (Jannin et al., 2006).

Two main aspects have been validated by the selected publications: the data acquisition process and the modelling phase. Validation data sets consist of fully simulated data from computers, data provided by simulated ORs, from phantoms, or real data directly from surgical interventions and patients. Computer simulations are one way of validating data that are easy to create, process, analyse and control, but are usually far from clinical reality. Similarly, virtual environments (simulated ORs) are also quite far from reality. While both approaches allow real flexibility for validating systems, it remains very difficult to model realistically a surgical environment, such as haptic feedbacks, anaesthesiological constraints, or surgeon/patient interaction. Moreover, even if the simulation is close to reality, the human factor is missing and could be an issue for applications that are intended to be used in real OR environments. The third possibility is to use real surgical devices on phantoms instead of humans. Even if the environment is closer to reality than complete virtual environments, it remains a part of the procedure that is not realistic. The validation strategies generally consist of, leave-one-out or k-fold cross-validation approaches. The comparison metrics are the recognition rate (accuracy), reproducibility, specificity and sensitivity.

Few evaluation studies have been conducted and reported in the literature (Katic et al., 2010; Ko et al., 2007; Yoshimitsu et al., 2010). Some papers indirectly showed the added value of the SPM approach through its use in comparing populations of surgical cases performed with different systems or by surgeons with different surgical expertise. For these few papers that evaluate their systems, the same possible limitations as the validation part can be expressed.

## 3. Similar works not included in the corpus

From the beginning of the 90s, many clinical studies were published which used the principle of time-motion analysis. Time was the first information chosen by teams to evaluate surgical systems, tools, approaches or to assess surgeons. Publications covering time-motion analysis are very close to the papers that are cited here from the data acquisition aspect. Indeed, they used off-line observer-based recording from videos (installed in the OR, on the surgeon, or in the operating field) for acquiring sequences of phases/steps/activities that are then processed through statistical analysis. The corresponding studies, mainly published in clinical journals, restricted their analysis to statistical computations of time or number of occurrences. They were not included in our corpus. Some major examples of publications are listed here: Weinger et al., 1994; den Boer et al., 2001; Sjoerdsma et al., 2000; Darzi and Mackay, 2002; Bann et al., 2003; Dosis et al., 2004; Mehta et al., 2002; Malik et al., 2003; Cao et al., 1999; Claus et al., 1995; Payandeh et al., 2002. A classification of their data acquisition techniques and modelling is proposed here in **Tab 2**.

	Data acquisition				Modelling		
	Granularity level	Operator +/- body part	Time of acquisition	Method for recording	Granularity level	Operator +/- body part	Formalisation
<b>Time-Motion analysis</b>	Steps/Activities/Motions	Surgeon	Intra	Observer-based recording from video (off-line)	Steps/Activities/Motions	Surgeon	Hierarchical decomposition

**Tab 2** - Classification of time-motion analysis publications, for the data acquisition and the modelling component.

Some recent papers used robot-supported recording, such as the paper of Hager et al. (2006) or Rosen et al. (2006). Fully connected HMMs were used for classifying hand trajectories to assess the level of surgeons' expertise. They were not included in our corpus since they did not incorporate any sequential aspect of the surgical processes. The models incorporated sequences of activities but these were not constrained. An existing recent review has already been published on the methods for objective surgical skills evaluation (Reiley et al., 2011), which includes all papers using trajectories analysis for surgical skills assessment. A non-exhaustive list of these papers is given here: Hager et al. (2006), Rosen et al. (2001, 2002, and 2006), Voros and Hager (2008), Lin et al. (2006). A classification of their data acquisition techniques and modelling is also proposed in **Tab 3**.

	Data acquisition				Modelling		
	Granularity level	Operator +/- body part	Time of acquisition	Method for recording	Granularity level	Operator +/- body part	Formalisation
<b>Surgical skill evaluation</b>	Motions	Surgeon	Intra-operative	Robot-supported recording (on-line)	Motions	Surgeon	Sequential list of words

**Tab 3** - Classification of surgical skills evaluation using robot-supported recording publications, for the data acquisition and the modelling component.

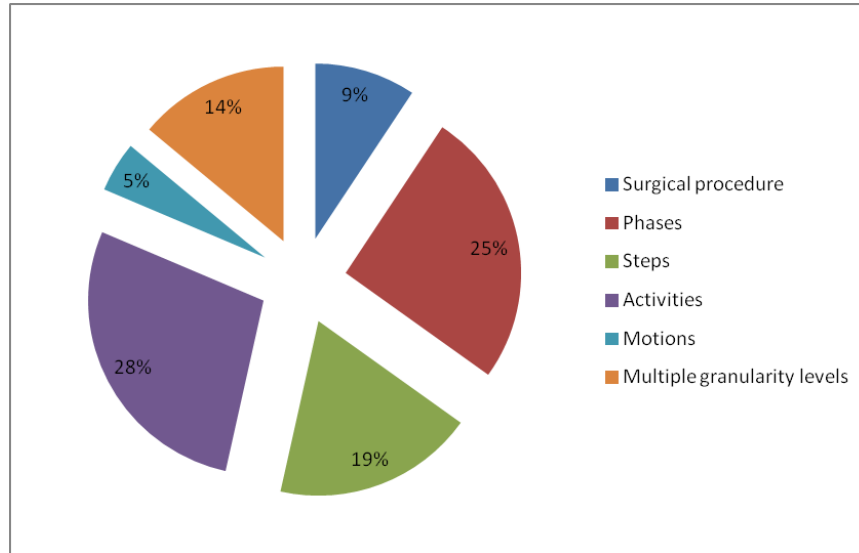
Others studies focused on the pre-processing steps before an SPM analysis. Radrich et al. (2008, 2009) presented a system for synchronising multi-modal information using various signals for surgical workflow analysis. Sielhorst et al. (2005) synchronised 3D movements before the comparison of surgeons' activities. Speidel et al. (2008, 2009) focused on the identification of instruments in MIS, with the goal of improving current intra-operative assistance systems.

With a methodology similar to SPM, some other studies focused on the modelling of the peri operative process, based on hospital systems (Wendt et al., 2003; Winter et al., 2003), hospital data (Maruster et al., 2001; Rosenbloom et al., 2006), or on surgical staff activities (Favela et al., 2007; Sanchez et al., 2008). Other research focused on the modelling of the OR environment (inside and outside) but without looking at the surgery itself (Riley and Manias, 2005; Sandberg et al., 2005; Archer and Macario, 2006). Their main objective is the improvement of the quality of patient care along with greater medical safety by studying flows or activities. Also, from an anaesthetist's point of view, work has been undertaken which looked at the ergonomics and organisation inside the OR (Seim et al., 2005; Schleppers and Bender, 2003; Decker and Bauer, 2003; Gehbard and Brinkmann, 2006). None of these studies focused on the surgical process and were therefore not included in our corpus.

## 4. Discussion

### 4.1 Modelling





**Fig 7** - Repartition (percentage of publications) of granularity levels of the modelling.

As we can see from **Fig 7**, all granularity levels have been studied, with a particular focus on steps and activities. Moreover, a consequent number of these studies use multiple granularity levels in their modelling. This type of approach seems to be required for creating global SPMs which integrate all aspects of the surgical procedure.

From the methods used for formalisation, XML schema, which is a lightweight ontology, defines a grammar that characterises the structure of a document or the type of data used. XML schemas can be a solution for describing SPMs at a high level of granularity, to structure data using a well-defined grammar, but they do not include important concepts such as classes or organisation into a hierarchy. In addition, they do not provide a relevant solution for representing the dynamic aspect of the process. As XML schema, the UML class diagram does not allow unique and uniform entities to be defined. Both approaches seem to be less suited to the formalisation of a surgical context than heavyweight ontologies. These allow two elements corresponding to the same unit to be specified. Unlike taxonomies that define classes and the relations between these classes, ontologies allow inference rules to be defined. Jannin et al. (2003) proposed a model based on the pre and post-operative acquisition of data, including interviews with surgeons. The types of surgical procedure, steps and actions were extracted and allowed the model to be created. Additionally, information related to images was linked to classes. Lemke et al. (2004) first defined a surgical ontology as a formal terminology for a hierarchy of concepts and their relationship in the specialised clinical context of surgical procedures and actions. Later, Burgert et al. (2006) proposed an explicit and formal description of an upper-level-ontology based on General Ontological Language (GOL) for representing surgical interventions. These pieces of work were the first to introduce heavyweight ontologies in the context of surgery.

Formalisation is crucial to be able to compare and share studies between different centres. Even though two centres acquire data about the same surgical procedure using the exact same terminology, a heavyweight ontology is still needed to be able to use both sets of data in a shared study, since this is the only way to ensure that a term has a single meaning in both studies. The more formalisation is used in the modelling, the more semantics will be considered and the more sharable the SPM will be. A heavy and rich formalisation is therefore the key for the future analysis of SPMs to tackle all these issues.

## 4.2 Data acquisition

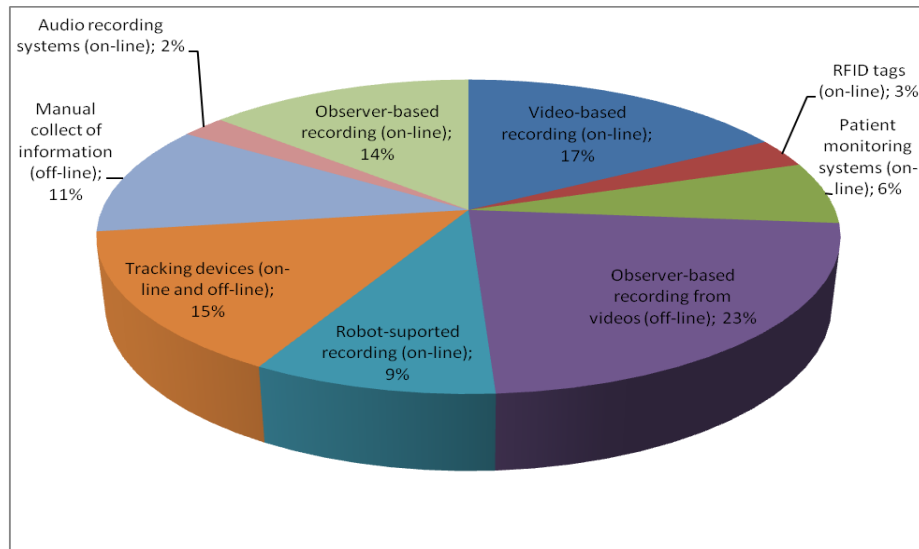
Both observer-based and sensor-based data acquisition approaches present advantages and drawbacks. Within observer-based approaches, the data acquisition process can be supported by two levels of knowledge: the description relies on *a priori* knowledge available thanks to common standards of surgical procedures or to fixed-protocol created by local experts. In the first case, standard surgical terms are reported for describing surgery whereas in the second case, the first

step consists of building up its own vocabulary. A new terminology is employed and permits a representation of knowledge that is particular to the surgeon's own experience and to the specific surgical environment. The related models are in most cases not based on an ontology and they are thus not an efficient formal representation of the knowledge and are also not easily sharable between centres. Moreover, the major concern of the on-line observer-based approach is the need for manual labelling that means that the system is not automatic, is time-consuming and prone to intra and inter observer variability. The necessity of having one person in the OR for recording, who is often an expert surgeon for reliable information labelling, is a strong inherent limitation to this approach. At the same time, it is the best way for recording finer details and capturing a high semantic level, which makes this technique advantageous compared to sensor-based approaches that do not acquire data at this level of semantics.

Sensor-based approaches are now increasingly adopted. For motion detection using tracking systems, the main drawback is that it relies on tools only and rare movements may not be efficiently detected due to the lack of dedicated training. Compared to other data acquisition techniques, analyses of videos would permit not only the installation of additional materials in the OR to be avoided, but also to have a source of information that does not have to be controlled by a human. For instance, acquiring information from the endoscopic view is very promising for high level information recognition. Videos are a very rich source of information, as demonstrated in laparoscopy by Speidel et al. (2008). Using image-based analysis, it is possible to acquire relevant information about surgery without disturbing the flow of the intervention. Unfortunately, current image-based algorithms, even with progress in computer vision, do not allow the well-known semantic gap to be captured in full, in which low-level visual features cannot correctly represent the high-level semantic content of images. For instrument use models, in spite of high detection accuracies, the major concern is that the recording of signals is not automatic when RFID tags are not used. The entire annotation is performed manually, which makes the system unusable in clinical routines. In practice, RFID tags are too intrusive, and some vital information that could improve the detection rates is missing, such as the anatomical structures treated. Generally speaking, all type of sensors additionally installed in the OR show promising results for the challenge of workflow recovery, but the main drawback is the modification of the OR set-up and the need to manage such new devices. In particular, eye-gaze tracking systems are interesting because they take into account the perceptual behaviour of the surgeon, but it would require large modifications during the intervention course not to alter the clinical routine as it stands.

In conclusion, observer-based approaches have the capacity to cover high granularity levels for describing surgery, from the lower level (time) to the highest, allowing the observer to take on the responsibility of acquiring semantic information from pre-defined terminologies and ontologies. On the other hand, it is a very time consuming and costly approach, with the need for a surgeon with a certain clinical background in the OR during the whole procedure. Sensor-based approaches do not have this ability to capture information with semantic meanings, but have the advantage of recording live signals automatically or semi-automatically, which is less time-consuming and allows the design of context aware systems.

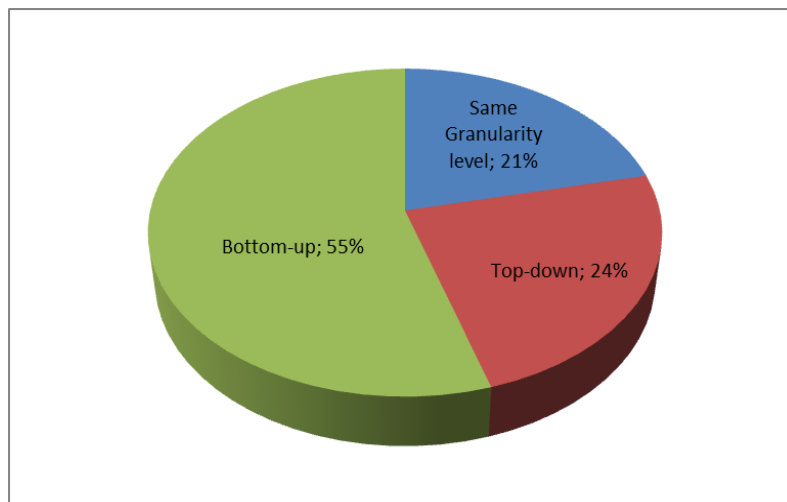
Currently, no papers cover multiple levels of granularity, which shows the difficulty of combining different data acquisition methods at different granularity levels. Multiple sensors can be used for instance for both capturing videos and the positions of instruments, but the combination of observer-based and sensor-based approaches turns out to be very difficult to set up. We see from **Fig 8** that no predominant techniques have been used.



**Fig 8** – Repartition (% of publications) of data acquisition techniques

### 4.3 Analysis

The choice of analysis methods that allow one to go from data to model is vital in SPM methodology. Bottom-up approaches are the most current (**Fig 9**). They allow a bridging of the semantic gap between numeric and symbolic data. Based on a preliminary formalisation, these methods all use supervised techniques based on a training stage, except for the work of Ahmadi et al. (2007). People report recognition rates of from 70% up to 99% but these values are very difficult to compare due to the differences of validation strategies as well as the differences in surgical specialities or the number and type of data used. The two others approaches (approaches that stay at the same granularity level and top-down approaches), even if they have still not completely demonstrated their interest for the field, are now more and more used.



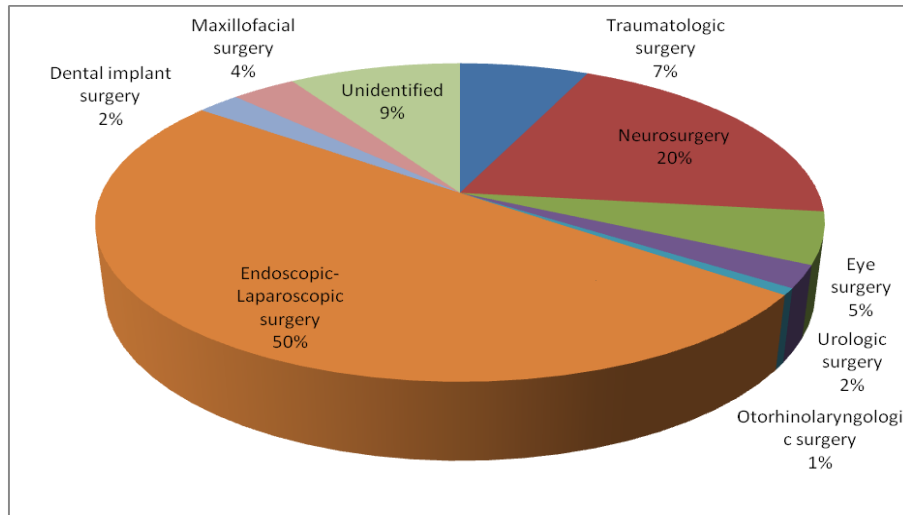
**Fig 9** - Repartition (% of publications) of the type of approaches used for “data to model” approaches.

The category of aggregation/fusion analysis method is important because it is a smart way of creating gSPMs that can be used as a supplementary tool for assisting surgeons. It allows creating procedural knowledge models based on an automated SPM analysis and not on traditional knowledge acquisition methods. The problem of this kind of approach is that it only represents the SPMs that are studied and may not cover all SP possibilities. Even if clearly it seems to be a vital aspect for improving surgical performance, no extensive work has been performed while this type of approach suggests good prospects in the future. Efforts must therefore be made here for integrating and automating average models of surgical processes in clinical routines.

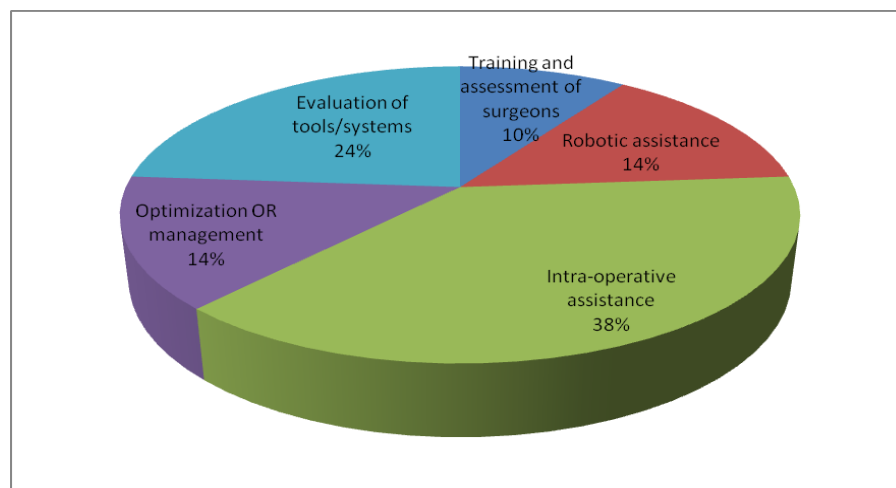
Similar to the previous category of the analysis approach, comparison and classification using surgical processes has not yet motivated many studies, but it may be a direction that needs to be considered. Comparisons of tool uses, surgeons or surgery performance using these kinds of methods allow a quantitative validation and assessment of the impacts on the surgical procedure.

#### 4.4 Applications

We have restricted in the diagram potential applications to the 5 most common ones cited in the papers. Additionally, when multiple applications were cited in the papers, we only used the main, clearly identified one. **Fig 10** shows the repartition of applications as well as the surgical specialities.



a)



b)

**Fig 10** - Repartition (% of publications) of surgical specialities (a) and clinical applications (b).

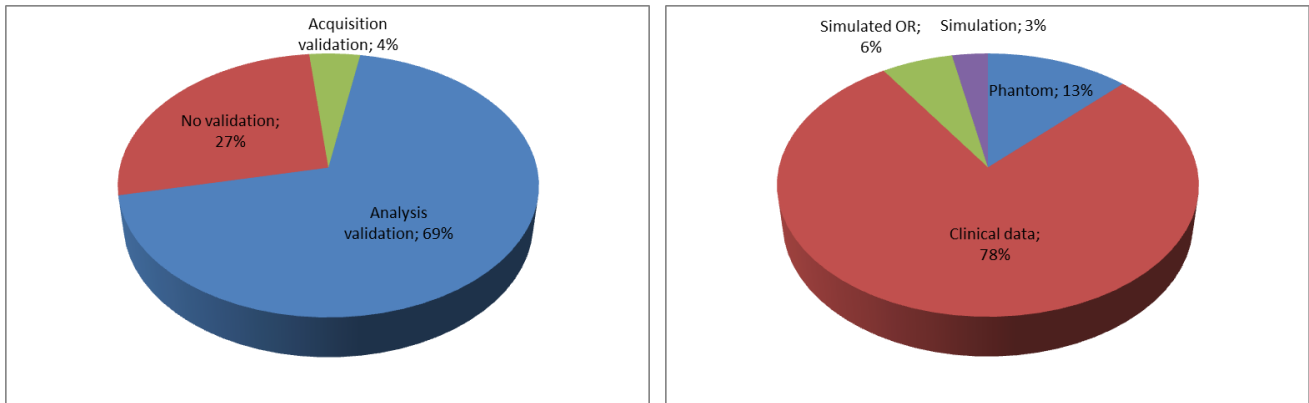
Most of the SPM studies were performed in the context of neurosurgery or endoscopy/laparoscopy. This is not surprising, as neurosurgery and MIS have been the most common applications used for computer-assisted surgery research. In the case of endoscopic and laparoscopic procedures, surgical procedures are often highly standardised, with a well-defined protocol, they are widely documented, and have inter-patient variability which remains very low. Data are also easily available for engineers for this surgical speciality. In neurosurgical procedures, data can also be easily acquired. In the case of eye surgery, new studies are using this surgical speciality because of the very short and

standardised procedures.

The distribution of applications is more uniform than the distribution of surgical specialties. Even if systems aiming at improving intra-operative assistance predominate, the four other applications have been seriously and similarly considered. Ahead of the large number of applications cited in publications, we see that SPMs can be useful along the entire surgery timeline, from pre-operative use to post-operative analysis. They can be used in every medical process and adapted to every surgical speciality, which shows the potential importance of SPMs.

#### 4.5 Validation-Evaluation

Most of the papers include validation studies (**Fig 11**, left) of the analysis part (69%), while only 4% of the papers validated the acquisition step. 27% of the papers do not validate their systems at all. When used, validations studies were performed (**Fig 11**, right) using clinical data in most cases (78%). Few studies use phantoms, simulated OR or computer simulations.



**Fig 11** - Repartition (% of publications) of the types of validation (left) and types of validation data (right).

Of the 46 publications that were peer-reviewed, only three of them performed evaluation studies. **Tab 4** shows the different elements of their evaluation studies.

Evaluation					
	System evaluated	Validation objective (Medical context)	Dataset	Metric	Operator
Katic et al. (2010)	Context-aware augmented reality system	Drilling planned implant	Phantom	Medical usability (questionnaire) Implant position comparison	Surgeon
Ko et al. (2007)	System for intelligent interaction scheme with a robot	Porcine Cholecystectomy	Clinical data	Number of voice commands	Surgeon
Yoshimitsu et al. (2010)	Scrub nurse robot	Endoscopic surgery	Clinical data	Instrument targeting time	Nurse

**Tab 4** - Classification of the 3 publications performing evaluation studies

However, no validation combined to evaluation has been conducted at the same time. This shows that research in the field, while being under considerable development, has not yet been introduced into the clinical routine.

#### 4.6 Correlations with other information

The correlation of SPMs with other information, such as patient-specific models, is an important prospect in the field. Patient-specific models are constructed from pre and post-operative patient data such as clinical data or images (Edwards et al., 1995; Biagioli et al., 2006; Verduijn et al., 2007; Kuhan et al., 2002). Being able to correlate patient

outcomes and pre-operative data with SPM would allow predictions to be made of the best possible surgical processes.

One other possibility would be to correlate SPMs with surgeons' decision-making processes during the intervention. The decision-making process in surgery can be conceptualised by two steps, the assessment and the diagnosis of the situation that must be used to select a specific action. The major aspect of the decision-making is that the decision depends on the level of expertise and tasks demanded. Dedicated models can be designed for surgical decision-making support by including this aspect. Moreover, correlation between pre and post-operative interviews of surgeons with the intra-operative intervention strategy would allow an analysis of surgeons' decision-making process to be made, especially under the pressure of time and a better understanding and anticipation of further adverse events (Flin et al., 2007; Jalote-Parmar et al., 2008; Morineau et al., 2009).

#### **4.7 Future of SPM**

Despite the potential impacts of SPM on computer assisted surgery outlined by the scientific and clinical communities, such methodology still needs to be deployed in clinical environments and applications and to demonstrate its added value. Some deadlocks remain. The first concerns the automatic acquisition and real time and robust monitoring of SPs. It seems clear that multi sensor approaches will be needed to reach high recognition rates at different granularity levels. Different points of view need to be used from closed sensors attached to operator's body, views of the operative field, signals from OR devices, patient's intraoperative data to large angle views of the whole operating room. Another issue relies on the computation of generic SPMs as the collection and gathering of possible SPs, as followed within an homogeneous population of surgical cases. Such generic SPMs constitute real procedural knowledge models (ref) and are needed to provide systems with a list of possible scenarios. However, they are limited by the data itself. Being sure that generic SPMs fully cover inter patient, inter surgeon, inter OR variability requires large worldwide data repositories with standardized terminologies and corresponding ontologies. The computation of generic SPMs also faces strong methodological issues in the aggregation/registration aspects, as a complex multi level sequence alignment problem. SPM methodology has the potential of allowing development of relevant comparison/classification approaches and metrics that could help understanding of surgical expertise. Where as the current developed metrics emphasize differences in practice, there is a need for methods explaining reasons of such differences. Finally, SPM methodology also needs to be seen by the clinicians as a skill augmentation support, a powerful teaching tool, rather than a "big brother" style-watching eye. Without a clear understanding of potential added value of the methodology by the clinicians, as well as a strong ethical awareness and control of the use of such data, such methodology will hardly be accepted by clinicians, increasing time from bench to bedside.

### **5. Conclusion**

Following the growing need for a new generation of CAS systems, new techniques have emerged based on the modelling of surgical processes. Research studies have been performed towards the development of sophisticated techniques for optimising, understanding and better managing surgeries and the OR environment based on SPMs. In this paper, we have presented a methodological review of the creation and the analysis of SPMs, focusing on works that modelled the procedural dimension. To organize the review, we have introduced a classification based on 5 major aspects of the SPM methodology: acquisition, modelling, analysis, application, and validation/evaluation. Using this classification, we have presented the existing literature and discussed the different existing methods and approaches. On the methodological side, we have shown that efforts still remain to be made in integrating the different granularity levels into a global framework. Both bottom-up and top-down approaches need to be combined. Methods are still needed to combine SPM into average generic SPMs. This methodological review has emphasised the possible large impact that SPM methodology may have in future surgical innovations as well as in surgical education, planning or intra-operative purposes. However, the technology is recent and there is still a lot of work to be done to demonstrate quantitatively its ethical added value within concrete clinical applications.

# References

- ✓ Agarwal S, Joshi A, Finin T, Yesha Y, Ganous T. A pervasive computing system for the operating room of the future. *Mobile Networks and Applications*. 2007; 12(2,3): 215-28.
- ✓ Ahmadi A, Sielhorst T, Stauder R, Horn M, Feussner H, Navab N. Recovery of surgical workflow without explicit models. *Med Image Comput Comput Assist Interv*. 2006; 9(1): 420-8.
- ✓ Ahmadi A, Padoy N, Rybachuk K, Feussner H, Heining SM, Navab, N. Motif discovery in OR sensor data with application to surgical workflow analysis and activity detection. M2CAI workshop, *Med Image Comput Comput Assist Interv*. 2009.
- ✓ Archer T and Macario A. The drive for operating room efficiency will increase quality of patient care. *Curr Opin Anaesthesiol*. 2006; 19: 171-6.
- ✓ Bann MS, Khan, Darzi A. Measurement of surgical dexterity using motion analysis of simple bench skills. *World J Surg*. 2003; 27: 390-4.
- ✓ Biagioli B, Scolletta S, Cevenini G, Barbini E, Giomarelli P, Barbini P. A multivariate Bayesian model for assessing morbidity after coronary artery surgery. *Crit Care*. 2006; 10(3): R94.
- ✓ Blum T, Padoy N, Feussner H, Navab N. Workflow mining for visualization and analysis of surgeries. *Int J Comput Assisted Radiol Surg*. 2008; 3(5): 379-86.
- ✓ Bouarfa L, Jonker PP, Dankelman J. Discovery of high-level tasks in the operating room. *J Biomed Inform*. 2011; 44(3): 455-62.
- ✓ Bhatia B, Oates T, Xiao Y, Hu P. Real-time identification of operating room state from video. *AAAI*. 2007; 1761-6.
- ✓ Burgert O, Neumuth T, Lempp F, Mudunuri R, Meixensberger J, Strauß G, Dietz A, Jannin P, Lemke HU. Linking top-level ontologies and surgical workflows. *Int J Comput Assisted Radiol Surg*. 2006; 1(1): 437-8.
- ✓ Cao CGL, MacKenzie CL, Payandeh S. Task and Motion Analysis in Endoscopic Surgery. *ASME Dynamic Systems, 5th Annual Symposium on Haptic Interface for Virtual Environment and Teleoperation*. 1996.
- ✓ Claus GP, Sjoerdsma W, Jansen A, Grimbergen CA. Quantitative standardised analysis of advanced laparoscopic surgical procedures. *Endosc Surg Allied Technol*. 1995; 3: 210-3.
- ✓ Cleary K, Chung HY, Mun SK. OR 2020: The operating room of the future. *Laparoendoscopic and Advanced Surgical Techniques*. 2005; 15(5): 495-500.
- ✓ Coles EC and Slavin G. An evaluation of automatic coding of surgical pathology reports. *J Clin Pathol*. 1976; 29(7): 621-6.
- ✓ Darzi A, Mackay S. Skills assessment of surgeons. *Surgery*. 2002; 131(2): 121-4.
- ✓ Decker K and Bauer M. Ergonomics in the Operating Room. *Minim Invasive Ther Allied Technol*. 2003; 12(6): 268-77.
- ✓ Den Boer KT, de Wit LT, Davids PHP, Dankelman J, Gouma DJ. Analysis of the quality and efficiency of learning laparoscopic skills. *Surg Endosc*. 2001; 15: 497-503.
- ✓ Dexter F, Epstein RH, Traub RD, Xiao Y. Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. *Anesthesiology*. 2004; 101(6): 1444-53.
- ✓ Dosis A, Bello F, Moorthy K, Munz Y, Gillies D, Darzi A. Real-time synchronization of kinematic and video data for the comprehensive assessment of surgical skills. *Stud Health Technol Inform*. 2004; 98:82-8
- ✓ Edwards FH, Peterson RF, Bridges C, Ceithaml EL. 1988: Use of a Bayesian statistical model for risk assessment in coronary artery surgery. Updated in 1995. *Ann Thorac Surg*. 1995; 59(6): 1611-2.
- ✓ Favela J, Tentori M, Castro LA, Gonzalez VM, Moran EB, Martinez-Garcia AI. Activity recognition for context-aware hospital applications: issues and opportunities for the deployment of pervasive networks. *Mobile Networks Applications*. 2007; 12(2,3): 155-71.
- ✓ Fischer M, Strauss G, Burgert O, Dietz A, Trantakis C, Meixensberger J, Lemke HU. ENT-surgical workflow as an instrument to assess the efficiency of technological developments in medicine. *Comput Assisted Radiol Surg*. 2005; 851-5.
- ✓ Flin R, Youngson G, Yule S. How do surgeons make intraoperative decisions. *Qual Saf Health Care*. 2007; 16: 235-9.
- ✓ Forestier G, Lalys F, Riffaud L, Trelhu B, Jannin P. Classification of surgical processes using dynamic time warping. *J Biomed Inform*. 2012; 45: 255-64.
- ✓ Gehbard F and Brinkmann, A. Management of an operating room in a university hospital. *Zentralbl Chir*. 2006; 131(4): 341-6.
- ✓ Hager G, Vagvolgyi B, Yuh D. Stereoscopic video overlay with deformable registration. *Medicine Meets Virtual Reality*. 2007.
- ✓ Haro BB, Zapella L, Vidal R. Surgical gesture classification from video data. *Med Image Comput Comput Assist Interv*. 2012; 7510: 34-41.
- ✓ Houliston BR, Parry DT, Merry AF. TADAA: Towards automated detection of anaesthetic activity. *Methods of Information in Medicine*. 2011; 50(5): 464-71.
- ✓ Hu P, Ho D, MacKenzie CF, Hu H, Martz D, Jacobs J, Voit R, Xiao Y. Advanced Visualization platform for surgical operating room coordination. Distributed video board system. *Surg innovation*. 2006; 13(2): 129-35.

- ✓ Ibbotson JA, MacKenzie CL, Cao CG, Lomax AJ. Gaze patterns in laparoscopic surgery. *Stud Health Technol Inform.* 1999; 7: 154-60.
- ✓ Jalote-Parmar A, van Alfen M, Hermans JJ. Workflow Driven User Interface for Radiological System: A Human Factors Approach. *Comput Assisted Radiol Surg.* 2008.
- ✓ James A, Vieira D, Lo BPL, Darzi A, Yang GZ. Eye-gaze driven surgical workflow segmentation. *Med Image Comput Comput Assist Interv.* 2007; 10(2):110-7.
- ✓ Jannin P, Raimbault M, Morandi X, Seigneuret E, Gibaud B. Design of a neurosurgical gestures model for multimodal image guided surgery. *Comput Assisted Radiol Surg.* 2001; 102-107.
- ✓ Jannin P, Raimbault M, Morandi X, Riffaud L, Gibaud B. Model of surgical procedures for multimodal image-guided neurosurgery. *Computer Aided Surgery.* 2003; 8(2): 98-106.
- ✓ Jannin P, Grova C, Maurer CR. Model for defining and reporting reference-based validation protocols in medical image processing. *Int J Comput Assist Interv.* 2006; 1(2): 1001-115.
- ✓ Jannin P Morandi X. Surgical models for computer-assisted neurosurgery. *Neuroimage.* 2007; 37(3): 783-91.
- ✓ Katic D, Sudra G, Speidel S, Castrillon-Oberndorfer G, Eggers G, Dillman R. Knowledge-based situation interpretation for context-aware augmented reality in dental implant surgery. *Med Imaging Augmented Reality.* 2010; 531-40.
- ✓ Klank U, Padoy N, Feussner H, Navab N. Automatic feature generation in endoscopic images. *Int J Comput Assisted Radiol Surg.* 2008; 3(3,4): 331-9.
- ✓ Ko SY, Kim J, Lee WJ, Kwon DS. Surgery task model for intelligent interaction between surgeon and laparoscopic assistant robot. *J Robotics Mechatronics.* 2007; 8(1): 38-46.
- ✓ Kuhan G, Marshall EC, Abidia AF, Chetter IC, McCollum PT. A Bayesian hierarchical approach to comparative audit for carotid surgery. *Eur J Vasc Endovasc Surg.* 2002; 24(6): 505-15.
- ✓ Kuhnappel U, Cakmak HK, Maab H. Endoscopic Surgery Training using Virtual Reality and Deformable Tissue Simulation. *Computer and Graphics.* 2000; 24: 671-82.
- ✓ Lalys F, Riffaud L, Bouget D, Jannin P. A framework for the recognition of high-level surgical tasks from video images for cataract surgeries. *IEEE Trans Biomed Eng.* 2012a; 59(4): 966-76.
- ✓ Lalys F, Bouget D, Riffaud L, Jannin P. Automatic knowledge-based recognition of low-level tasks in ophthalmological procedures. *Int J Comput Assist Radiol Surg.* 2012b; 8(1): 39-49.
- ✓ Lemke HU, Trantakis C, Köchy K, Müller A, Strauss G, Meixensberger J. Workflow analysis for mechatronic and imaging assistance in head surgery. *Int Congress Series.* 2004; 1268: 830-5.
- ✓ Lin HC, Shafran I, Yuh D, Hager GD. Towards automatic skill evaluation: Detection and segmentation of robot-assisted surgical motions. *Computer Aided Surgery.* 2006; 11(5): 220-30.
- ✓ Lo B, Darzi A, Yang G. Episode Classification for the Analysis of Tissue-Instrument Interaction with Multiple Visual Cues. *Med Image Comput Comput Assist Interv.* 2003; 1: 231-7.
- ✓ MacKenzie CL, Ibbotson AJ, Cao CGL, Lomax A. Hierarchical decomposition of laparoscopic surgery: a human factors approach to investigating the operating room environment. *Minim Invasive Ther Allied Technol.* 2001; 10(3): 121-8.
- ✓ Malik R, White P, Macewen C. Using human reliability analysis to detect surgical error in endoscopic DCR surgery. *Clin Otolaryngol Allied Sci.* 2003; 28: 456-60.
- ✓ Maruster L, van der Aalst W, Weijters T, van den Bosch A, Daelemans W. Automatic discovery of workflows models from hospital data. *BNAIC.* 2001; 183-90.
- ✓ Marvik R, Lango T, Yavuz Y. An experimental operating room project for advanced laparoscopic surgery. *Semin Laparosc Surg.* 2004; 11: 211-6.
- ✓ Mehta NY, Haluck RS, Frecker MI, Snyder AJ. Sequence and task analysis of instrument use in common laparoscopic procedures. *Surgical Endoscopy.* 2002; 16(2): 280-5.
- ✓ Meng F, D'Avolio LW, Chen AA, Taira RK, Kangarloo H. Generating models of surgical procedures using UMLS concepts and multiple sequence alignment. *Am Med Inform Assoc Annu Symp Proc.* 2005; 520-4.
- ✓ Meyer MA, Levine WC, Egan MT, Cohen BJ, Spitz G, Garcia P, Chueh H, Sandberg WS. A computerized perioperative data integration and display system. *Int J Comput Assisted Radiol Surg.* 2007; 2(3,4): 191-202.
- ✓ Miyawaki F, Masamune K, Suzuki S, Yoshimitsu K, Vain J. Scrub nurse and timed-automata-based model for surgery. *IEEE Industrial Electronics Trans.* 2005; 5(52): 1227-35.
- ✓ Morineau T, Morandi X, Le Moëllic N, Diabira S, Haegelen C, Hénaux PL, Jannin P. Decision making during preoperative surgical planning. *Human factors.* 2009; 51(1): 66-77.
- ✓ Münchenberg J, Brief J, Raczkowsky J, Wörn H, Hassfeld S, Mühling J. Operation Planning of Robot Supported Surgical Interventions. *Int Conf Intelligent Robots Systems.* 2000; 547-52.
- ✓ Nara A, Izumi K, Iseki H, Suzuki T, Nambu K, Sakurai Y. Surgical workflow monitoring based on trajectory data mining. *New frontiers in Artificial Intelligence.* 2011; 6797: 283-91.
- ✓ Neumuth T, Schumann S, Strauss G, Jannin P, Meixensberger J, Dietz A, Lemke HU, Burgert O. Visualization options for surgical workflows. *Int J Comput Assisted Radiol Surg.* 2006a; 1(1): 438-40.
- ✓ Neumuth T, Durstewitz N, Fischer M, Strauss G, Dietz A, Meixensberger J, Jannin P, Cleary K, Lemke HU, Burgert O. Structured recording of intraoperative surgical workflows. *SPIE medical imaging - PACS in Surgery.* 2006b: 6145; 61450A.



- ✓ Neumuth T, Trantakis C, Eckhardt F, Dengl M, Meixensberger J, Burgert O. Supporting the analysis of intervention courses with surgical process models on the example of fourteen microsurgical lumbar discectomies. *Int J Comput Assisted Radiol Surg.* 2007; 2(1): 436-8.
- ✓ Neumuth T, Jannin P, Strauss G, Meixensberger J, Burgert O. Validation of Knowledge Acquisition for Surgical Process Models. *J Am Med Inform Assoc.* 2008; 16(1): 72-82.
- ✓ Neumuth T, Loebe F, Jannin P. Similarity metrics for surgical process models. *Artif Intell Med.* 2012a; 54(1): 15-27.
- ✓ Neumuth T, Liebmann P, Wiedemann P, Meixensberger J. Surgical workflow management schemata for cataract procedures. Process Model-based design and validation of workflow schemata. *Methods Inf Med.* 2012b; 51(4).
- ✓ Neumuth T and Meissner C. Online recognition of surgical instruments by information fusion. *Int J Comput Assisted Radiol Surg.* 2012c (*In press*)
- ✓ Nomm S, Petlenkov E, Vain J, Belikov J, Miyawaki F, Yoshimitsu K. Recognition of the surgeon's motions during endoscopic operation by statistics based algorithm and neural networks based ANARX models. *Proc Int Fed Automatic Control.* 2008.
- ✓ Padoy N, Horn M, Feussner H, Berger M, Navab N. Recovery of surgical workflow: a model-based approach. *Int J Comput Assisted Radiol Surg.* 2007; 2(1): 481-2.
- ✓ Padoy N, Blum T, Feuner H, Berger MO, Navab N. On-line recognition of surgical activity for monitoring in the operating room. *Conf Inno App Art Intell.* 2008.
- ✓ Padoy N, Blum T, Ahmadi SA, Feussner H, Berger MO, Navab N. Statistical modeling and recognition of surgical workflow. *Med Image Anal.* 2010; 16(3): 632-41.
- ✓ Payandeh S, Lomax AJ, Dill J, Mackenzie CL, Cao CGL. On Defining Metrics for Assessing Laparoscopic Surgical Skills in a Virtual Training Environment. *Stud Health Technol Inform.* 2002; 85:334-40.
- ✓ Patel, V. L., J. F. Arocha, et al. A primer on aspects of cognition for medical informatics. *J Am Med Inform Assoc.* 2001; 8(4): 324-43.
- ✓ Payne PRO, Mendonca EA, Johnson SB, Starren JB. Conceptual knowledge acquisition in biomedicine: a methodological review. *J Biomed Inform.* 2007; 40(5): 582-602.
- ✓ Qi J, Jiang Z, Zhang G, Miao R, Su Q. A surgical management information system driven by workflow. *IEEE Conf Service Operat Logist Inf.* 2006; 1014-8.
- ✓ Radrich H. Vision-based motion monitoring trough data fusion from a chirurgical multi-camera recording system. Diploma thesis. TUM, Munich. 2008.
- ✓ Radrich H, Padoy N, Ahmadi A, Feussner H, Hager G, Burschka D, Knoll A. Synchronized multimodal recording system for laparoscopic minimally invasive surgeries. *M2CAI workshop, Med Image Comput Comput Interv.* 2009.
- ✓ Rattner WD, Park A. Advanced devices for the operating room of the future. *Seminars in laparoscopic surgery.* 2003; 10(2): 85-8.
- ✓ Reiley CE, Lin HC, Yuh DD, Hager GD. Review of methods for objective surgical skill evaluation. *Surg Endosc.* 2011. 25(2), 356-66.
- ✓ Riedl S. Modern operating room mangement in the workflow of surgery. Spectrum of tasks and challenges of the future. *Der Chirurg.* 2002; 73: 105-10.
- ✓ Riley R and Manias E. Governing Time in Operating Rooms. *J Clin Nurs;* 2005;15(5); 546-53.
- ✓ Rosen J, Hannaford B, Sinanan M, Solazzo M. Objective Evaluation Of Laparoscopic Surgical Skills Using Hidden Markov Models Based On Haptic Information And Tool/tissue Interactions. *Stud Health Technol Inform.* 2001; 81: 417-23.
- ✓ Rosen J, Solazzo M, Hannaford B, Sinanan M. Task decomposition of laparoscopic surgery for objective evaluation of surgical residents' learning curve using hidden markov model. *Comput Aided Surg.* 2002; 7(1): 49-61.
- ✓ Rosen J, Brown JD, Chang L, Sinanan M, Hannaford B. Generalized Approach for Modeling Minimally Invasive Surgery as a Stochastic Process Using a Discrete Markov Model. *IEEE Trans Biomed Eng.* 2006; 53(1): 399-413.
- ✓ Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Facilitating direct entry of clinical data into electronic health record systems. *J Am Med Inform Assoc.* 2006; 13(3): 277-88.
- ✓ Sanchez D., Tentori, M., Favela, J. Activity recognition for the smart hospital. *IEEE intelligent systems.* 2008; 23(2): 50-7.
- ✓ Sandberg WS, Daily B, Egan MT, Stahl JE, Goldman JM, Wiklund RA, Rattner D. Deliberate perioperative systems design improves operating room throughput. *Anesthesiology.* 2005; 103: 406-18.
- ✓ Satava RM and Carrico CJ. Advanced Simulation Technologies for Surgical Education. *Bulletin of the American College of Surgeon.* 1996; 81(7): 71-7.
- ✓ Satava R, Cuschieri A, Hamdorf J. Metrics for objective assessment. *Surg Endosc.* 2003; 17(2):220-6.
- ✓ Schleppers A and Bender H. Optimised workflow and organisation – from the point of view of an anaesthesiolo-g department. *Minim Invasive Ther Allied Technol.* 2003; 12(6): 278-83.
- ✓ Seim AR, Meyer M, Sandberg WS. Does parallel workflow impact anaesthesia quality. *Am Med Inform Assoc Annu Symp Proc.* 2005; 1053.

- ✓ Sielhorst, T., Blum, T., Navab, N. Synchronizing 3d movements for quantitative comparison and simultaneous visualization of actions. *Int Symp Mix Augm Real*. 2005; 38-47.
- ✓ Sjoerdsma W, Meijer D, Jansen A, den Boer KT, Grimbergen CA. Comparison of efficiencies of three techniques for colon surgery. *J Laparoendosc Adv Surg Tech*. 2000; 10(1): 47-53.
- ✓ Speidel S, Sudra G, Senemaud J, Dretnschew M, Müller-Stich BP, Gun C, Dillmann R. Recognition of risk situations based on endoscopic instrument tracking and knowledge based situation modeling. *Progress Biomed Optics Imaging*. 2008; 9(1): 35.
- ✓ Speidel S, Benzko J, Krappe S, Sudra G, Azad P, Müller-Stich BP, Gutt C, Dillmann R. Automatic classification of minimally invasive instruments based on endoscopic image sequences. *Progress Biomed Optics Imaging*. 2009; 10(1): 37.
- ✓ Sudra G, Speidel S, Fritz D, Möller-Stich BP, Gutt C, Dillmann R. MEDIASSIST: MEDICAL ASSITance for intraoperative skill transfer in minimally invasive surgery using augmented reality. *Progress Biomed Optics Imaging*. 2007; 8(2).
- ✓ Suzuki T, Yoshimitsu K, Tamura M, Muragaki Y, Iseki H. Video information management system for information guided neurosurgery. *Comp Aided Surg*. 2012; 3(3), 75-82.
- ✓ Taylor C, Draney MT, Ku JP, Parker D, Steele BN, Wang K, Zarins CK. Predictive medicine: computational techniques in therapeutic decision making. *Comp Aid Surg*. 1999; 4: 231-47.
- ✓ Thienjarus S, James A, Yang GZ. An eye-hand data fusion framework for pervasive sensing of surgical activities. *Pattern Recognition*. 2012; 45(8): 2855-67.
- ✓ Verduijn M, Rosseel PM, Peek N, de Jonge E, de Mol BA. Prognostic Bayesian networks II: an application in the domain of cardiac surgery. *J Biomed Inform*. 2007; 40(6): 619-49.
- ✓ Voros S and Hager GD. Towards “real-time” tool-tissue interaction detection in robotically assisted laparoscopy. *Int Conf IEEE Biomed Robot Biomechat*. 2008: 562-7.
- ✓ Weinger MB, Herndon OW, Zornow MH, Paulus MP, Gaba DM, Dallen LT. An objective methodology for task analysis and workload assessment in anesthesia providers. *Anesthesiology*. 1994; 80(1); 77-92.
- ✓ Wendt T, Häber A, Brigl B, Winter A. Modeling hospital information systems (Part 2): using the 3LMG2 tool for modelling patient, record management. *Methods Inf Med*; 2003; 43(3); 256-67.
- ✓ WFMC – Workflow management coalition. Terminology & glossary. Doc number WFMC-T-1011, Issue 3.0. Winchester, UK.1999.
- ✓ Winter A, Brigl B, Wendt T. Modeling Hospital information systems (Part 1): The revised three-layer graph-based meta model 3LGM2. *Method Inf Med*. 2003; 42(5): 544-51.
- ✓ Xiao, Y., Hu, P., Hu, H., Ho, D., Dexter, F., Mackenzie, CF., Seagull, FJ. An algorithm for processing vital sign monitoring data to remotely identify operating room occupancy in real-time. *Anesth Analg*, 2005; 101(3): 823-32.
- ✓ Yoshimitsu K, Masamune K, Iseki H, Fukui Y, Hashimoto D, Miyawaki F. Development of scrub nurse robot (SNR) systems for endoscopic and laparoscopic surgery. *Micro-NanoMechatronics Human Science*. 2010; 83-88.
- ✓ Yule S, Flin R, Paterson-Brown S, Maran N. Non-technical skills for surgeons in the operating room: a review of the literature. *Surgery*. 2006; 193(2): 140-9.