

# **Supporting Pathology Process Management with Real-time Business Intelligence**

**Wei Chen Li**

Thesis submitted to the  
Faculty of Graduate and Postdoctoral Studies  
in partial fulfillment of the requirements for the degree of

**Master of Computer Science**

Under the auspices of the Ottawa-Carleton Institute for Computer Science



**uOttawa**

University of Ottawa  
Ottawa, Ontario, Canada

October 2016

© Wei Chen Li, Ottawa, Canada, 2016

# Abstract

---

[Context] Clinical pathology is a medical specialty that uses laboratory analysis of tissues and fluids to diagnose diseases. As emerging methodologies and technologies extend their realm to anatomical pathology, pathology facilities are faced with new challenges related to the increasing volume and complexity of patient cases in their work processes. [Problem] Existing tools that monitor process data show limitations in supporting pathology process management. They lack the capability of identifying emerging process bottlenecks in a real-time manner, and this prevents facility managers from conducting corrective actions proactively. [Methodology] Applying the Design Science Research Methodology, this thesis proposes and builds a Business Intelligence solution that provides visual analysis and real-time monitoring functions to support pathology process management. A usability study with expert participants from the Eastern Ontario Regional Laboratory Association was conducted to validate this solution. [Results] The proposed solution meets the requirements of the experts and can provide efficient, flexible, and multi-tenancy support to users in the context of their daily professional activities.

# Acknowledgment

---

First and foremost, I would like to express my deepest appreciation to my supervisors, Professor Daniel Amyot, Professor Lysanne Lessard, and Professor Wojtek Michalowski for their untiring support and motivation. I am feeling so grateful for their patience and trust in me. This work would not have been possible without their guidance and assistance. Not only have they been sharing their invaluable knowledge with me, but they have also been teaching me good attitude and manners on how to view, analyze, and resolve problems in academic and everyday life.

I would like to thank my manager from IBM, Graham Fasken, for his continuous support and understanding. I would not have been able to complete my studies on time without his guidance on work-life balance and support on flexible working hours.

I would also like to thank my colleague and technical mentor, Dr. Randy Giffen, for leading me into this study domain and for sharing his valuable knowledge and experience.

Sincere thanks to Dr. Diponkar Banerjee, Dr. Fawaz Halwani, Mr. William Park, and the Department of Pathology and Laboratory Medicine at The Ottawa Hospital, for continuously supporting the study, sharing knowledge on the management of pathology processes, and helping with the usability study.

Special thanks to my colleagues, Alvin Cho and Rob Kelly, for sharing insights and always encouraging me. Thanks for being friends and big brothers.

Last but not least, no words can express how grateful I am towards my family members. They have sacrificed a lot to support me working and studying at the same time for years. Thanks for always being with me and supporting me.

# Table of Contents

---

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgment.....</b>	<b>iii</b>
<b>Table of Contents .....</b>	<b>iv</b>
<b>List of Figures.....</b>	<b>vii</b>
<b>List of Tables .....</b>	<b>viii</b>
<b>List of Acronyms .....</b>	<b>ix</b>
<b>Chapter 1. Introduction .....</b>	<b>1</b>
1.1. Motivation.....	1
1.2. Research Problem and Research Objective.....	2
1.3. Methodology .....	3
1.4. Thesis Contributions .....	4
1.5. Publications .....	4
1.6. Thesis Outline .....	4
<b>Chapter 2. Background.....</b>	<b>6</b>
2.1. Background on Business Intelligence.....	6
2.1.1 Business Intelligence Systems.....	6
2.1.2 BI Solution Business Context.....	7
2.1.3 BI Solution Abstract Architecture .....	7
2.1.4 BI Dashboard.....	9
2.1.5 Dashboard Data Visualization .....	11
2.1.6 OLTP vs. OLAP .....	13
2.2. Background on Pathology Work Process .....	14
2.2.1 Terminology .....	14
2.2.2 Pathology Process Overview .....	16
2.3. Chapter Summary .....	18
<b>Chapter 3. Business Context and Requirements .....</b>	<b>19</b>
3.1. Business Drivers .....	19
3.2. Business Goals.....	19
3.3. Chapter Summary .....	22

<b>Chapter 4. Literature Review .....</b>	<b>23</b>
4.1. <i>Overall Usage of BI in Healthcare .....</i>	23
4.2. <i>Keywords, Queries, and Study Collection .....</i>	25
4.3. <i>Study Categorization, Comparison and Analysis .....</i>	26
4.3.1 <i>Overview and Categorization .....</i>	26
4.3.2 <i>Comparison and Analysis .....</i>	28
4.4. <i>Chapter Summary .....</i>	30
<b>Chapter 5. Architecture .....</b>	<b>31</b>
5.1. <i>Abstract Architecture of the Solution.....</i>	31
5.2. <i>Data Tier.....</i>	33
5.3. <i>Application Tier .....</i>	34
5.4. <i>Presentation Tier .....</i>	35
5.5. <i>Client Tier.....</i>	36
5.6. <i>Chapter Summary .....</i>	36
<b>Chapter 6. Development and Demonstration.....</b>	<b>37</b>
6.1. <i>Tools Selection.....</i>	37
6.2. <i>Data Tier Development.....</i>	38
6.2.1 <i>Restored Database .....</i>	38
6.2.2 <i>Additional Data Tables .....</i>	40
6.3. <i>Application Tier Development .....</i>	40
6.3.1 <i>Data Model Implementation .....</i>	40
6.3.2 <i>Calculation Formulas.....</i>	42
6.3.3 <i>Aggregation, Sorting, and Parameters .....</i>	43
6.4. <i>Presentation Tier Implementation .....</i>	45
6.4.1 <i>Dashboard Layout Design .....</i>	45
6.4.2 <i>Pending Case Analysis Component.....</i>	46
6.4.3 <i>Pathologist Workload Analysis Component.....</i>	50
6.4.4 <i>Case Count Analysis Component .....</i>	52
6.4.5 <i>Case Detail Report.....</i>	53
6.5. <i>Client Tier Implementation .....</i>	54
6.6. <i>Enabling Real-time Monitoring in the Dashboard.....</i>	56
6.7. <i>Chapter Summary .....</i>	57
<b>Chapter 7. Evaluation .....</b>	<b>58</b>
7.1. <i>Objective .....</i>	58
7.2. <i>Methods.....</i>	59
7.3. <i>Participant Recruitment.....</i>	60

7.4.	<i>Participation</i> .....	61
7.4.1	Overall Description.....	61
7.4.2	Predefined User Scenario .....	62
7.5.	<i>Task Result Analysis</i> .....	63
7.5.1	Timing Analysis .....	63
7.5.2	Correctness Analysis .....	66
7.5.3	Multi-dimensional Analysis.....	68
7.6.	<i>Questionnaire</i> .....	70
7.7.	<i>Chapter Summary</i> .....	74
<b>Chapter 8.</b>	<b>Discussion</b> .....	<b>75</b>
8.1.	<i>Technical Comparison</i> .....	75
8.2.	<i>User Feedback on Existing Features</i> .....	77
8.2.1	User Satisfaction about the Stakeholder Requirement Fulfillment.....	77
8.2.2	User Feedback on Decision Making.....	78
8.3.	<i>Limitations and Threats to Validity</i> .....	78
8.4.	<i>Chapter Summary</i> .....	80
<b>Chapter 9.</b>	<b>Conclusions and Future Work</b> .....	<b>81</b>
9.1.	<i>Contributions</i> .....	81
9.2.	<i>Future Work</i> .....	82
<b>References</b> .....		<b>84</b>
<b>Appendix A:</b>	<b>Contact Email Sent to Potential Participants</b> .....	<b>88</b>
<b>Appendix B:</b>	<b>Consent Letter</b> .....	<b>89</b>
<b>Appendix C:</b>	<b>Experiment Protocol</b> .....	<b>92</b>
<b>Appendix D:</b>	<b>Participant Tasks</b> .....	<b>94</b>
<b>Appendix E:</b>	<b>Task Answers and Instructions</b> .....	<b>96</b>
<b>Appendix F:</b>	<b>Questionnaire</b> .....	<b>99</b>
<b>Appendix G:</b>	<b>Data Tier Database Structure</b> .....	<b>103</b>
<b>Appendix H:</b>	<b>Data Model Logical Layer Diagram</b> .....	<b>107</b>

# List of Figures

---

<b>Figure 1</b>	DSRM process model (adapted from [42]) .....	3
<b>Figure 2</b>	BI dashboard abstract architect diagram (adapted from [18]) .....	9
<b>Figure 3</b>	Block with tissue fragment embedded in paraffin .....	15
<b>Figure 4</b>	Block being sliced before staining and slide production <sup>4</sup> .....	15
<b>Figure 5</b>	Tray with 20 slides where different stains were used <sup>4</sup> .....	16
<b>Figure 6</b>	Pathology process overview .....	17
<b>Figure 7</b>	Literature review process .....	23
<b>Figure 8</b>	BI in healthcare operation and performance management loop (adapted from [35]) .....	24
<b>Figure 9</b>	EORLA BI project abstract architect diagram .....	32
<b>Figure 10</b>	Dashboard layout .....	46
<b>Figure 11</b>	Pending Case Analysis upper section design .....	47
<b>Figure 12</b>	Pending Case Analysis lower section overview .....	49
<b>Figure 13</b>	Pending Case Analysis value prompt options for drop-down lists .....	49
<b>Figure 14</b>	Pathology Workload Analysis first version .....	50
<b>Figure 15</b>	Pathology Workload Analysis final version .....	52
<b>Figure 16</b>	Case Count Analysis .....	53
<b>Figure 17</b>	The case detail report .....	53
<b>Figure 18</b>	Prompt page of stand-alone case detail report .....	54
<b>Figure 19</b>	Pathology dashboard showed on a laptop computer .....	55
<b>Figure 20</b>	Pathology dashboard showed on an Apple iPad .....	55
<b>Figure 21</b>	Predefined user scenario .....	63
<b>Figure 22</b>	Task time measurement for all participants (minimum, average, maximum) .....	64
<b>Figure 23</b>	Total spent time in minutes, sorted in increasing order .....	65
<b>Figure 24</b>	Pending Case Analysis correctness pie charts .....	67
<b>Figure 25</b>	Pathologist Workload Analysis correctness pie charts .....	67
<b>Figure 26</b>	Case Count Analysis correctness pie chart .....	68
<b>Figure 27</b>	Cross analysis by participant scatter chart .....	69
<b>Figure 28</b>	Questionnaire results on the “use and learn” experience .....	70
<b>Figure 29</b>	Questionnaire results on UI presentation .....	71
<b>Figure 30</b>	Questionnaire results on user interaction .....	72
<b>Figure 31</b>	Questionnaire results on advanced features .....	72
<b>Figure 32</b>	Total satisfaction score radar chart .....	73
<b>Figure 33</b>	Questionnaire results on management role .....	74
<b>Figure 34</b>	Data model logical layer diagram (application tier) .....	107

# List of Tables

---

<b>Table 1</b>	MAD framework.....	10
<b>Table 2</b>	Dashboard type mapping.....	11
<b>Table 3</b>	Chart type suggestions for data features.....	12
<b>Table 4</b>	OLTP vs. OLAP comparison.....	13
<b>Table 5</b>	EORLA BI project requirements .....	21
<b>Table 6</b>	Summary of the literature review of the pathology BI dashboard.....	26
<b>Table 7</b>	List of selected tools.....	37
<b>Table 8</b>	Sample data files characteristics .....	38
<b>Table 9</b>	Data model implantation work summary .....	41
<b>Table 10</b>	Aggregation, sorting and parameters in the dashboard components ..	44
<b>Table 11</b>	“Heat map-like” colour palette for Pathology Workload Analysis ....	51
<b>Table 12</b>	Multi-dimensional analysis by task result table.....	70
<b>Table 13</b>	Comparison of the thesis with the closely-related work .....	77

# List of Acronyms

---

<b>Acronym</b>	<b>Definition</b>
BI	Business Intelligence
CRAP	Contrast, Repetition, Alignment and Proximity Design Principle
CSV	Comma-Separated Value
DPLM	Department of Pathology and Laboratory Medicine
DSRM	Design Science Research methodology
DSS	Decision Support System
EORLA	Eastern Ontario Regional Laboratory Association
ETL	Extract, Transform, and Load
IOM	Institute of Medicine
JDBC	Java Database Connectivity
KPI	Key Performance Indicator
MAD	Monitor, Analyze, and Detail
OLAP	On-line Analytical Processing
OLTP	On-line Transaction Processing
SQL	Structured Query Language
TOH	The Ottawa Hospital
UI	User Interface

# Chapter 1. Introduction

---

Clinical pathology is a medical specialty that uses laboratory analysis of tissues and fluids to diagnose diseases. This thesis uses Business Intelligence technologies to develop a new real-time dashboard for managing pathology processes and operations. The usability and potential usefulness of this dashboard is validated by a group of operational managers of pathology processes and facilities. This chapter highlights the thesis motivation, research question, methodology, and main contributions.

## 1.1. Motivation

With the rapid development of science and technology, new methodologies and technologies extend their realm to anatomical pathology. Hospitals have adopted new technologies and emerging personalized medicine that ask for targeted pathological diagnoses in the anatomical pathology practice. In this context, important challenges are being faced such as the increasing volume of patient cases and complexity of pathology work processes. In order to overcome these challenges, not only are pathologists required to use efficient ways to diagnose patient cases, but extra work and improvements are required by the organization (department and laboratory) and operational managers of pathology facilities. If the pathology management could identify emerging process bottlenecks in real time, conduct corrective actions, and then assess their results, they would be able to ensure that the whole pathology work process is efficient and that all pathology specimens are processed on time.

Business Intelligence (BI) technologies are widely adopted in different industries, including healthcare, and much evidence confirms the ability of BI to assist organization management. Given the challenges being experienced in pathology departments, historical data about case processing and laboratory performance is no longer sufficient to manage them effectively. As existing tools cannot satisfy the operational performance requirements posed by pathology departments and the increasing demand from pathology management,

BI technologies have much potential to help a pathology organization determine process bottlenecks, maintain good quality processes, and maximize effectiveness.

The need for supporting pathology process management can be addressed by providing a *real-time* monitoring BI dashboard, that is, with data refreshed as specimens are examined according to pathology process stages, in laboratories. With the emergence of dashboards as tools for integration of key performance indicators (KPIs), real-time monitoring automatically presents relevant and summarized information in a timely fashion. Management of pathology facilities could use such solution to monitor the working status of their departments, take early actions when a stage of the process is bottlenecked, and improve the quality of care by optimizing the process from case accession to diagnostic reporting.

## **1.2. Research Problem and Research Objective**

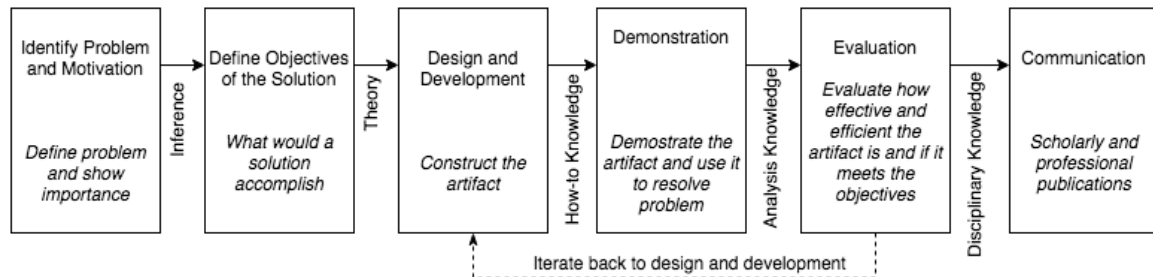
The main research problem of this thesis is “*use of historical data for managing pathology facility does not satisfy current management needs*” and the research objective is “*a real-time business intelligence solution exploiting barcode tracking data can provide efficient, flexible, and multi-tenancy support to pathology department management*”.

To further define the scope of the research objective, we clarify the concepts that it references:

- *Real-time*: the real-time characteristic of the solution refers to the capability of a dashboard to auto-refresh at a user-defined rate that is fast enough to support informed tracking (for example, by refreshing results every minute).
- *Flexible*: the flexibility characteristic of the solution refers to the feature through which users have easy access to the dashboard, from different types of device, e.g., laptop or mobile devices, and have the capability of interacting with the dashboard and customizing the dashboard components.
- *Multi-tenancy*: multi-tenancy refers to the capacity of the dashboard to serve users with different roles and deliver the right information for each target user group.

### 1.3. Methodology

The methodology adopted in this thesis research is the Design Science Research Methodology (DSRM) proposed by Peffers et al. [42] and illustrated in Figure 1.



**Figure 1** DSRM process model (adapted from [42])

To conduct the thesis research according to the DSRM Process Model, the following steps were performed in the study:

1. Identify research problem and motivation: identify challenges in pathology facilities for the management of high volumes and complex work processes;
2. Define the objective of the solution: requirements for a real-time BI solution to support real-time pathology process management;
3. Design and development: define the architecture, design, and develop a BI prototype satisfying the stakeholder requirements;
4. Demonstration: demonstrate the proposed BI solution as a proof of concept in order to show how to use it to resolve problems. Exploit data collected in real-time (through barcode scanning of specimens tracked during the pathology process) in a large pathology facility;
5. Evaluation: perform a usability study (with tasks and questionnaires) of a limited number of expert participants to evaluate the proposed solution using different measures;
6. Communication: publish peer-reviewed articles and write this thesis to describe the research problem and results.

## 1.4. Thesis Contributions

To resolve the research problem and accomplish the research objective discussed in Section 1.2, a fully functional BI solution prototype was design and implemented, and a usability study involving a limited number of expert participants (operations managers of different levels) was conducted.

The major research contributions of this thesis study include, (1) further confirmed and demonstrated that using BI solution is an efficient way to help improve an organization work process management; (2) gathered a set of generalized stakeholder requirements that can be used in any pathology management BI solution design and implementation; (3) proposed a BI dashboard layout design with a combination of colour indictors, tables and diagrams for the pathology workload management; and (4) proposed a low-cost method to enable real-time monitoring in BI dashboard.

The major practical contributions of this thesis study include the implementation of a four-tier BI architect solution, and the usability study to evaluated dashboard.

## 1.5. Publications

This thesis resulted in two publications so far. The first one focuses on initial prototype developed by Dr. Halwani and preliminary results on the design of the new real-time dashboard presented in Chapter 6. The second publication proposes extensions of this thesis work towards predictive analytics.

1. Halwani, F., **Li, W.C.**, Banarjee, D., Lessard, L., Amyot, D., Michalowski, W., and Giffen, R.: “A real-time dashboard for managing pathology processes”. *Journal of Pathology Informatics*, Vol. 7, Article 24, Wolters Kluwer (2016)
2. Lessard, L., Michalowski, W., **Li, W.C.**, Amyot, D., Halwani, F., and Banarjee, D.: “Predictive Analytics to Support Real-Time Management in Pathology Facilities”. *AMIA 2016 Annual Symposium Proceedings*, Chicago, USA, November. AMIA Knowledge Center (2016, to appear)

## 1.6. Thesis Outline

The thesis is organized as follows.

- The thesis continues in Chapter 2 with background information about the BI domain and the pathology work process.
- Chapter 3 introduces the business context of the case study (Eastern Ontario Regional Laboratory Association – EORLA<sup>1</sup>) and discusses the first stage of the DSRM process.
- Chapter 4 presents the literature review for existing closely-related work and industry solutions for the research problem, and targets on finding out the research gap.
- Chapter 5 describes the abstract architecture of the BI solution, which aims to guide the design and development of the thesis work independently of the implementation technologies. This partly covers the development and demonstration stage in the DSRM process
- Chapter 6 presents the dashboard design and implementation of the EORLA pathology BI solution, hence also covering the development and demonstration stage in the DSRM process.
- Chapter 7 describes the usability study performed in order to evaluate the proposed BI solution, for the evaluation stage of the DSRM process.
- Chapter 8 compares the proposed tool with closely related work against the stakeholder requirements and discusses user feedback about the tool. Threats to the validity of the thesis work are also discussed in this chapter.
- Finally, Chapter 9 concludes the thesis and recommends future work items.

Note that Appendices A to F describe the content of letters and forms used for the usability study in Chapter 7, while Appendices G and H provide additional precisions on database models used in Chapter 6.

---

<sup>1</sup> <http://www.eorla.ca/>

## Chapter 2. Background

---

This chapter introduces the background information of the thesis study. It starts with major definitions and concepts in the business intelligence domain, and then introduces the pathology related terminology and work process.

### 2.1. Background on Business Intelligence

In 1989, Dresner, from the Gartner Group, promoted “business intelligence (BI)” as an umbrella term, which describes “a set of concepts and methods to improve business decision making by using fact-based support systems” [44]. In the following subsections, the BI concepts and methods related to this thesis are introduced.

#### 2.1.1 Business Intelligence Systems

In the IBM Journal (1958), Luhn described a business intelligence system as a system that can auto-abstract information from different resources, auto-encode the extracted data, and auto-create and update for each action-point profile for an organization [33]. Since then, the concept of business intelligence system has been gradually developed and extended, and became a hot topic in both academic research and industry. In 2008, Negash and Gray defined a BI system as a data-driven decision support system that supports “data gathering, data storage, and knowledge management with analysis to provide input to the decision process” [40].

A BI system is capable of transforming data into usable, actionable information for various business purposes. To adopt a BI system, an organization requires the:

- Collection of sufficient and quality data for the business purpose;
- Installation and deployment of the BI techniques and applications;
- Motivation and skills to develop a BI application and apply it to the business.

### 2.1.2 BI Solution Business Context

The business context of a BI solution is established by: the business drivers, the goals, and the strategies. These elements define the fundamentals of a BI solution, which is “the business processes to be affected, the kinds of business applications to be implemented, and the information services to be provided” [47]. From a business measurement perspective, these three elements “describe why to measure, what to measure, and where to measure” [47].

- *Business drivers* determine the reasons why the BI solution is required and motivate the needs to take actions; they facilitate the identification of the user cases. Common BI solution drivers include: business competition, bottlenecks in workforce performance, and business market shrinking.
- *Business goals* determine the desired achievement of what needs to be accomplished according to the drivers; they help define the measurable results. The goals of a BI solution that responds to the driver examples could be: win the business competition, improve the workforce performance, and boost the business market.
- *Business strategies* describe what actions need to take to achieve the business goals; they make it easier to determine the business processes in need. The strategies that respond to the above examples could be: pricing competitively, providing sufficient support and training to the employees, and establishing new sales channels.

### 2.1.3 BI Solution Abstract Architecture

BI system architecture contains four tiers: data tier, application tier, presentation tier, and client tier. While the data tier and application tier are grouped to the back end, the latter two are usually considered as the front end, as illustrated in Figure 2.

The *data tier* is sometimes referred to as the foundation tier. As the name “foundation” suggests, this tier sits at the bottom of the architecture and provides the required databases for the BI applications to be running on. The data tier is composed of databases with three types of information, which are commonly categorized as content store, data sources, and business rules. The *content store*, also called application database, is usually

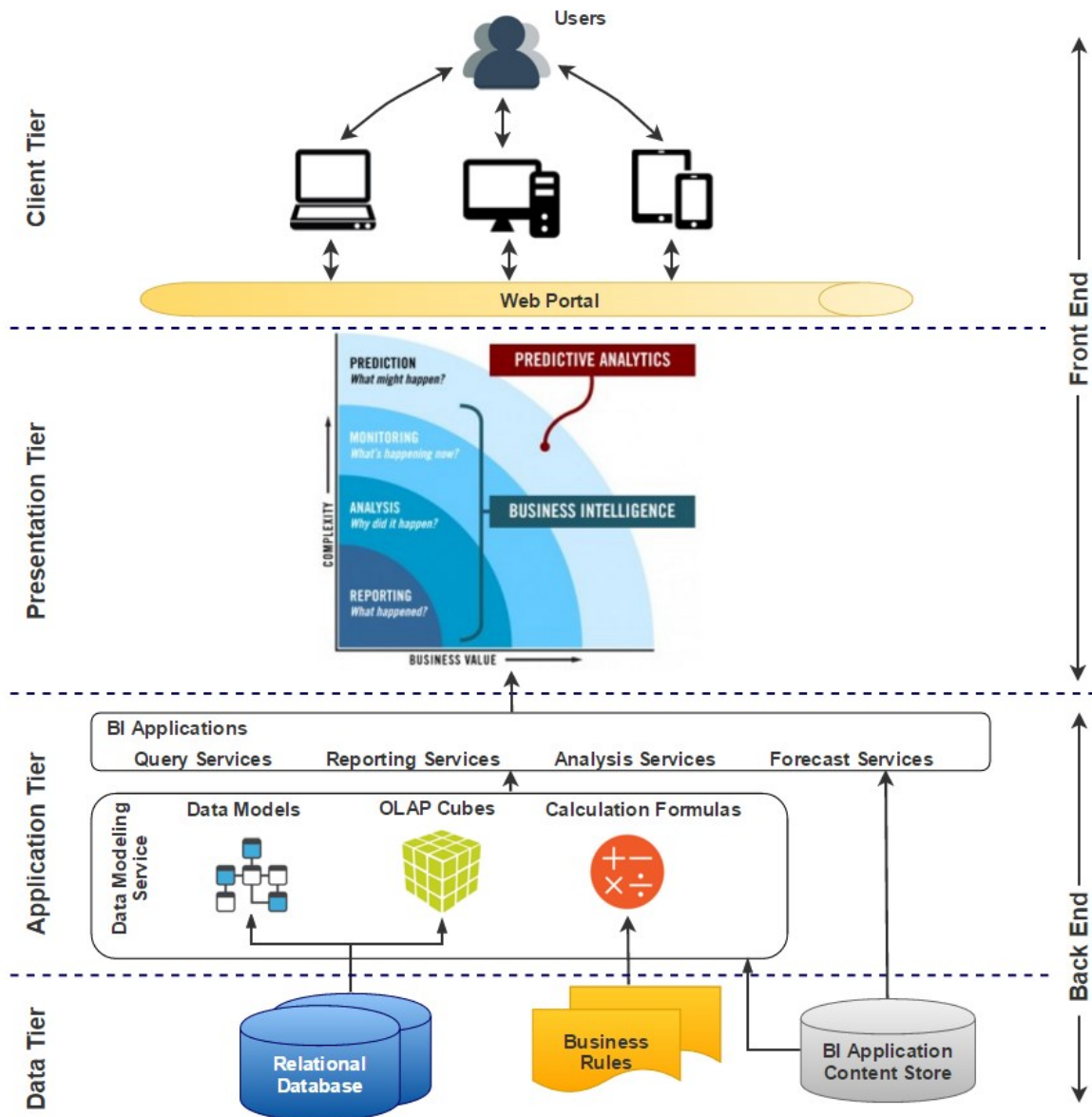
a relational database that supports the operation of BI applications. The *data source* is a relational database that, in case of this research, stores pathology department operation records. It keeps historical data of the department business operations and accepts new records in real time while the operation systems are generating new transactions. Data stores in the data sources are the resources that need to be transformed and analyzed, and are the sources against which the BI applications run. *Business rules* define what algorithms and aggregations need to apply to the transactions stored in the data source. Other than stored in a structured database, the business rules are sometimes saved in a flat file, as it is easier for a business manager to describe business rules in natural language. In addition to data storage, the data tier also provides connection information and methods for enabling BI applications to communicate with the databases, for example, using Java Database Connectivity (JDBC) connections.

The *application tier* is the core center that controls the BI mission. It handles different kinds of incoming requests and distributes the corresponding service to each request, for instance query services, reporting services, and analysis services. These types of services are usually pre-built into BI tools. Some advanced tools like IBM® Watson™ Analytics<sup>2</sup> also provide additional features such as predictive analysis. At the application tier, BI developers are not required to build the services because they are embedded into the tool. However, they need to invest some effort in developing the data modeling part. The data modeling related work usually includes: (1) building data models, where one constructs the logical data models from the data source physical data model; (2) building OLAP cubes, which are multi-dimensional data cubes that enable business analysis functionalities such as slicing and dicing; and (3) building calculation formulas, which translate business rules from natural language to mathematical equation or formulas that are understandable by the BI services.

BI solutions are presented to the users at the *client tier*. Some solutions are built with desktop-based user interfaces, while a web-based user interface is an alternative choice.

---

<sup>2</sup> IBM Watson Analytics, <https://www.ibm.com/analytics/watson-analytics/us-en/>



**Figure 2** BI dashboard abstract architect diagram (adapted from [18])

### 2.1.4 BI Dashboard

The term dashboard was firstly used in the automobile industry to indicate section of a vehicle where drivers can monitor the major status of a vehicle at a glance. A key performance indicator (KPI) is a performance measurement that evaluates the achievement or success of the key activities of an organization against its goals, and usually is considered to reason about strategic objectives. A business intelligence dashboard, by extending the concept of dashboard, provides an at-a-glance view of the KPIs of an organization to the

users, who usually come from the management of the organization. BI dashboards typically represent the past and current KPIs using charts and tables, and because of the need of showing the current data, dashboards are not designed to be static, but either real-time or refreshed periodically, for example, hourly, daily, or monthly. In the book *Information Dashboard Design*, Few defined a dashboard as “a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance” [12].

In Eckerson’s *Performance Dashboard* [8], the author suggested that a dashboard is constructed with one or more of three functionalities: monitoring, analysis, and management. He also introduces the MAD (Monitor, Analyze, and Detail) framework (Table 1), which maps three layers of information data types to the corresponding users. In this framework, data become more aggregated and concentrated if the layer number is smaller, while data volume and the number of users increases if the layer number is larger.

**Table 1** MAD framework

	<b>Functionality</b>	<b>Information Type</b>	<b>User Type</b>
<b>Layer 1</b>	High-level Monitor	Graphical Data	Executives
<b>Layer 2</b>	Analyze	Dimensional Data	Analysts
<b>Layer 3</b>	Detail	Transactional Data	Workers

Following that framework, three types of dashboards have been defined and they each emphasize a functionality (Table 2). The operational dashboard helps the front-line workers and managers monitor the operational activities, emphasize event-based the monitoring feature. The tactical dashboard is used by managers to monitor and analyze the department performance, putting more weight on analysis. The strategic dashboard enables executives and high level managers to track the process performance against the strategic target, and is for the management function. It was found that some application systems might integrate the characteristics and functions of multiple dashboard types, while other systems might be specifically dedicated to one type (Table 2).

**Table 2** Dashboard type mapping

<b>Functionality</b>	<b>Dashboard Type</b>	<b>User Type</b>
Management	Strategic Dashboard	Executives/Board
Analysis	Tactical Dashboard	Managers/Analysts
Monitoring	Operational Dashboard	Operations staff

### 2.1.5 Dashboard Data Visualization

Tufte in *The Visual Display of Quantitative Information* made the definition of data graphics or data visualization as “visually display measured quantities by means of the combined use of points, lines a coordinate system, numbers, symbols words, shading and colour” [48]. In BI dashboards, data visualization is being widely used because it is capable of consolidating and aggregating data into a collective visual context. Instead of leaving the users being buried under raw data, data visualization quickly captures one’s attention and exposes data trends to the users efficiently.

Charts help dashboard users to visualize data. However, sometimes it is a real challenge for a dashboard developer to find the right data visualization to present the data and meet user requirements, because there are too many choices of chart types. There is research on data graphical guideline, including [21][26][53], and a recent categorization that is widely used is proposed by Abela [1], as summarized in Table 3.

There are two types of data visualizations based on how a dashboard developer includes a chart in a dashboard. In a type-based approach, users can create one of a pre-defined fixed set of charts (bar chart, line chart, pie chart, etc.) and customize the chart using an Application Programming Interface (API) to set details such as colours and axis labels. In a language-based approach, users need to write a human-readable text language to describe a chart by a composable set of features. For example, a pie chart is a composition of: (1) an interval element (which will make the pie slices); (2) polar transform (to transform the intervals into slices); (3) stacking operation (to place the slices on top of each other); (4) colour and labeling aesthetics (to colour and label the slices). As new data visualizations being continuously developed and type-based visualizations are not capable of

providing all the chart types, for example, variable width column chart in Table 3, language-based data visualizations are gradually becoming popular in both academic research [51] and industry [7][19].

**Table 3** Chart type suggestions for data features

<b>Purpose</b>	<b>Data Feature</b>	<b>Chart Type</b>
Comparison, among items	Two variables per item	Variable width column chart
	One variable per item, many categories	Table or table with embedded charts
	One variable per item, few categories, many items	Bar chart
	One variable per item, few categories, few items	Column chart
Comparison, over time	Many periods, cyclical data	Circular area chart
	Many periods, non-cyclical data	Line chart
	Few periods, single or few categories	Column chart
	Few periods, many categories	Line chart
Composition, changing overtime	Many periods, only relative difference matter	Stacked 100% area chart
	Many periods, relative and absolute difference matter	Stacked area chart
	Few periods, only relative difference matter	Stacked 100% column chart
	Few periods, relative and absolute difference matter	Stacked column chart
Composition, static	Simple share of total	Pie chart
	Accumulation or subtraction to total	Waterfall chart
	Components of components	Stacked 100% column chart with subcomponents
Distribution	Single variable, few data points	Column histogram
	Single variable, many data points	Line histogram
	Two variables	Scatter chart
	Three variables	3D area chart
Relationship	Two variables	Scatter chart
	Three variables	Bubble chart

### 2.1.6 OLTP vs. OLAP

The *On-line Transaction Processing* (OLTP) is a class of information systems designed to support the business process nature of an organization. The requirement drives the system to be capable of handling a large number of short on-line transactions, for example inserting new operational records, in a way that is highly efficient in retrieving a small set of historical records. Database and data modeling in OLTP are very fast on data injection and query processing and needs low maintenance on data integrity.

As its name implies, the *On-line Analytical Processing* (OLAP), on the other hand, is a class of information systems focused on the analytical purpose of an organization. OLAP is widely used to perform complex queries for Decision Support Systems (DSS) or data mining, and thus has pre-calculated aggregations built in a multi-dimensional data structure like star or snowflake schema in data warehouses or during data modeling.

Some major differences between OLTP and OLAP [41] are listed in Table 4.

**Table 4** OLTP vs. OLAP comparison

	<b>OLTP</b>	<b>OLAP</b>
Data source	Operational data	Consolidated data from OLTP
What is stored	Snapshot of the structured on-going business process transactional data	Pre-summarized aggregations in multi-dimensional data cubes
Usage	Fast store transactional data, quick data retrieval	Analysis, decision support, and data mining
Insert and update	Fast and can be done in real-time	Long running, usually done by off peak hour batch jobs
Queries	Relatively simple queries	Complex queries with aggregations
Processing performance	High performance	Relatively slow

## 2.2. Background on Pathology Work Process

Clinical pathology is a medical specialty that uses laboratory analysis of tissues and fluids to diagnose diseases. This section introduces basic terminology and gives an overview of a typical pathology work process.

### 2.2.1 Terminology

Pathology includes several main sub-specialties, including:

- *Autopsy*: for deceased patients.
- *Cytology*: for single cells (smear, fluid, brush, discharge, aspirate)
- *Surgical*: for tissues or organs (biopsy, excision, resection). Specimens do not come necessarily from surgery; they can be medical (endoscopy), from a clinic (dermatology), or from a radiology department.

A *specimen* is a tissue to be assessed and diagnosed. In typical pathology laboratory:

1. Autopsy and Cytology have only one specimen per case (defined below).
2. Surgical has 1 to many (up to 50) specimens per case; they may be all from the same organ (e.g., different areas from the same breast), different organs that can be diagnosed by pathologist with the same sub-specialty (esophagus, stomach, duodenum), or occasionally different sub-specialties (e.g., heart + thymus + lymph node).
3. Each specimen is identified by the case number (as above) followed by a sequential number, i.e., SP-15-001-01, SP-15-001-02.
4. Each specimen requires its own assessment.

A *case* is defined as one or several specimens taken from the same patient and analyzed by the same pathologist, unless these specimens require assessment by pathologists with different main sub-specialties. Cases are identified uniquely. For example, in the Sunquest Information System's PowerPath®<sup>3</sup> tracking system:

---

<sup>3</sup> <http://www.sunquestinfo.com/products-solutions/anatomic-pathology>

1. Each case is given a unique number, prefixed by a specialty/subspecialty identifier, followed by the year received (YY), followed by a sequential number that resets back to 1 at the beginning of each year.
2. Each case is also given a unique report template that is also useful to identify the main case type.
3. Surgical sub-specialties do not have any identifiers, but they can be inferred from specimen codes and specimen descriptions using natural language pattern matching with about 90% accuracy.

*Blocks* are small (2 x 2 x 0.5 cm) rectangular boxes that isolate specimen fragments during processing, embedding in paraffin (Figure 3) and cutting to generate one to multiple slides per block (Figure 4).



**Figure 3** Block with tissue fragment embedded in paraffin<sup>4</sup>



**Figure 4** Block being sliced before staining and slide production<sup>4</sup>

---

<sup>4</sup> Ed Uthman, <https://www.flickr.com/photos/euthman/sets/72157594338538591/>, used with his permission

*Slides* are thin pieces of a block. Different stains are used to highlight different suspected diseases. Haemotoxylin and Eosin (H&E) is the most common stain, but depending on the disease, a pathologist may order additional slides from the same block/s stained differently for bugs (special stains), antigens (immunohistochemistry) or molecular studies. There are about 250 different stains available at the academic pathology lab.

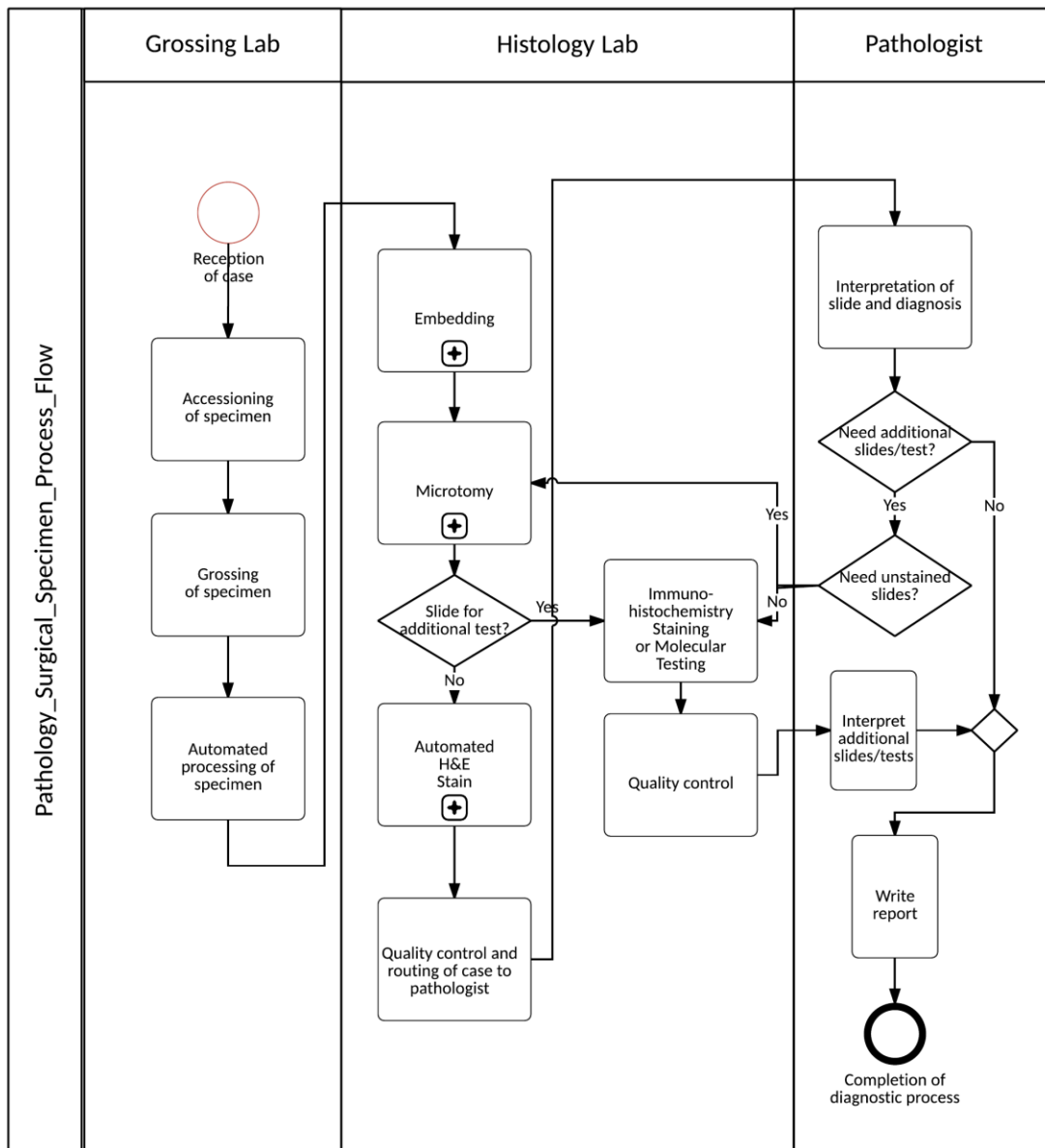


**Figure 5** Tray with 20 slides where different stains were used<sup>4</sup>

Each case has unique identifier that is subsequently specialized for specimens, blocks, and slides; in that sense, a case represents a master identifier while specimens are children. Case represent a unit of interest in terms of what is happening in the pathology workflow.

## 2.2.2 Pathology Process Overview

A typical pathology process is composed of a series of activities, or stages, as illustrated in Figure 6 (process activities are grouped by labs and stages).



**Figure 6** Pathology process overview

**Accessioning of specimen**

Containers with the specimens and paper clinical orders from physicians are first recorded while accessioning specimens. Case information is entered to the system; a unique ID case code is generated and labels are printed. Information for each specimen associated with a case is entered here manually and each specimen is associated with a pathology code for specification of how many blocks and slides need to be created for this specimen.

### **Grossing and automated processing**

This is where tissue is cut into blocks and these blocks are automatically processed so they can be later used for creating the slides.

### **Histology Lab**

Ready blocks are processed here. Each block is manually sliced to create the required number of slides. Slides are stained and verified.

### **Quality control and routing**

This is where finished slides are organized in the trays (per case and specimen); they are double-checked for correct matching with specimen and case codes; slide barcodes are scanned here and this scan indicates the end of the preparation process and signals that case is ready for dispatch and diagnosis by a pathologist. Each case is assigned to a pathologist; cases are batched and wait for being transported to pathologists' offices.

### **Pathologist's diagnosis**

Slides organized in a tray arrive in the pathologist's office. The pathologist receives a complete case (which may include multiple trays). When the pathologist starts a diagnostic process, he/she scans one slide (this indicates that case is being opened for a diagnosis) and proceeds with a diagnosis. During the process, the pathologist may order new/more slides or blocks. Orders of slides are captured electronically while orders of blocks are not (verbal communication only). An electronic order puts a pause in a diagnostic process until the order is received. Making a definite diagnosis completes the process.

## **2.3. Chapter Summary**

In this chapter, we discussed important background for the thesis study. We started with definitions and concepts in the business intelligence domain (including dashboard types, data visualization, and a generic BI architecture), and then introduced the pathology work process terminology and phases, illustrated with EORLA's process.

The next chapter introduces the business context and requirements of this thesis' pathology BI project. These requirements will be used as evaluation criteria in the literature review of Chapter 4 and in the dashboard discussion of Chapter 8.

## Chapter 3. Business Context and Requirements

---

To address the research problem and achieve the research objective discussed in Section 1.2, we conducted a case study for the EORLA (The Eastern Ontario Regional Laboratory Association) BI project. This chapter introduces the business drivers and business goals of this project.

### 3.1. Business Drivers

The Eastern Ontario Regional Laboratory Association is "a member-owned, non-profit organization encompassing the operation of 19 licensed, acute-care, hospital-based clinical laboratories that service clinical programs across the Champlain Local Health Integration Network (CLHIN) of Eastern Ontario" [9]. It currently amalgamates sixteen hospitals, with all surgical specimens being processed in one central location (Department of Pathology and Laboratory Medicine of The Ottawa Hospital) that houses grossing, histology, and cytology laboratories. A rapid growth and influx of surgical and cytology specimens have created many challenges in insuring smooth workflow and timely case processing by management and lab personnel. Although the entire pathology process is tracked by a barcode scanning system (PowerPath® by Sunquest) and operational data is stored in a local clinical database, the current solution (PowerJ [16]) lacks support for data reporting, data analysis, and pre-emptive warnings or alarms that can help the management to act proactively when an issue or performance bottleneck arises.

### 3.2. Business Goals

In general, in order to manage pathology workflows effectively, it is not enough to only rely on tracking historical data about case accessions, lab operations, and pathologist workloads. Managers need to be able to (1) identify emerging workflow bottlenecks in real-time, (2) conduct corrective actions, and (3) assess the results of their actions.

To quantitatively convert the conceptual business goals into measurable goals, we elicited requirements of the EORLA pathology BI project (Table 5). In this requirement list, several are standard BI dashboard requirements, e.g., *Tabular Data* and *Data Visualization*, while others are advanced features gathered by interviewing stakeholders of the project (EORLA managers and researchers). The requirements list was validated and agreed to by EORLA managers, and published in [16].

Requirement 1, *Tabular Data*, is a common requirement for dashboards as it provides a type of report with detailed and structured data than could be presented in a graphical format. Requirement 2, *Data Visualization*, is core to dashboard capabilities and helps pathology managers to quickly grasp the state of their facility and the presence of potential issues. Requirement 3, *Colour Indicators*, refers to the use of colour coding to indicate performance status. A well-recognized colour scheme is that of traffic lights, where green indicates that a target is being met (for example, a specimen is being processed well within a standard time frame), red indicates that a target is not being met (for example, a specimen has spent more hours in the histology lab than what is expected), and yellow indicates that the element being measured is approaching an undesired state (for example, a specimen has already spent 80% of the accepted time in the histology lab). Requirements 1 to 3 were identified as the user interface visualization requirements of the dashboard.

Requirements 4 to 7 were identified for the dashboard's user interactions. Requirement 4, *Value Filtering*, allows users to focus on a subset of data, for example only specimens that have been dispatched to pathologists for a diagnosis. Requirement 5, *Sorting*, enables users of the dashboard to sort data according to varied criteria, for example by date. The ability to *Drill Through* to event-level data for each case and each specimen (requirement 6) was a key motivator for the development of this thesis' dashboard since it would allow managers to pinpoint the exact source of delays within the pathology process. Requirement 7, *Interface Customization*, is a common stakeholder requirement for dashboards. Indeed, it is recognized that varied types of visualizations communicate information differently, for example, line graphs to track trends over time vs. bar charts to compare a set of values. Being able to customize the type of display helps to align users' information needs with the most adequate type of visual display (especially under constrained screen size).

**Table 5** EORLA BI project requirements

Requirement			Description
User Interface Visualization	1	Tabular Data	Displaying summary or detailed data in a list or pivot table.
	2	Data Visualization	Displaying summary data in a chart or diagram.
	3	Colour Indicators	Conveying performance status through a meaningful colour scheme.
User Interaction	4	Value Filtering	Dynamic filtering of report output by user selection.
	5	Sorting	Dynamic sorting of report output by user setting.
	6	Drill Through	Drilling through to event-level data for each case and each specimen
	7	Interface Customization	Switching between chart view/table views, or changing to different chart types.
Advanced Features	8	Real-time Reporting	Automatic updating of data analysis in real-time (in minutes or seconds).
	9	Stakeholder-oriented Views	Tailoring of the dashboard for varied user groups and roles.
	10	Specimen-type Customization	Specifying different business rules for different types of specimens.
	11	Automatic Notifications	Sending alerts to users by email or other channels.
	12	Workload Planning	Predicting future workloads based on historical data using prediction models.

Requirement 8, *Real-time Reporting*, was initially stated as the need for daily reporting of cases, slides, and specimen status. However, through the use of the PowerJ prototype by EORLA, it evolved to the need for real-time reporting to identify delays and backlog more rapidly. Real-time here means automatic updates configurable to a refresh rate fast enough to support informed tracking and decision-making (e.g., in the order of minutes or even seconds). Requirement 9, *Stakeholder-oriented Views*, helps to adapt the information presented to the needs and purpose of each stakeholder group and ensure that the right information is presented to the right group, in line with their decision-making needs. Requirement 10, *Specimen-type Customization*, incorporates different business rules as per specimen type in the case performance calculation algorithm, and enables the dashboard to

measure the status of the cases by specimen type level. Requirement 11, *Automatic Notifications*, stems from the desire of managers to be alerted of issues proactively by e-mail or other channels rather than having to login to a system to check the status of the pathology facility. Finally, Requirement 12, *Workload Planning*, forecasts future job duration using prediction models and assists managers with workflow planning to avoid bottlenecks. Requirements 8 to 12 were advanced features requested by the stakeholders for the dashboard after using PowerJ.

### 3.3. Chapter Summary

This chapter identified a major problem faced by organizations such as EORLA and motivates the need for a solution that the one currently available. Twelve generic requirements for a better solution were identified and explained. Note that several of these requirements are directly in line with the research problem concepts defined in Section 1.2:

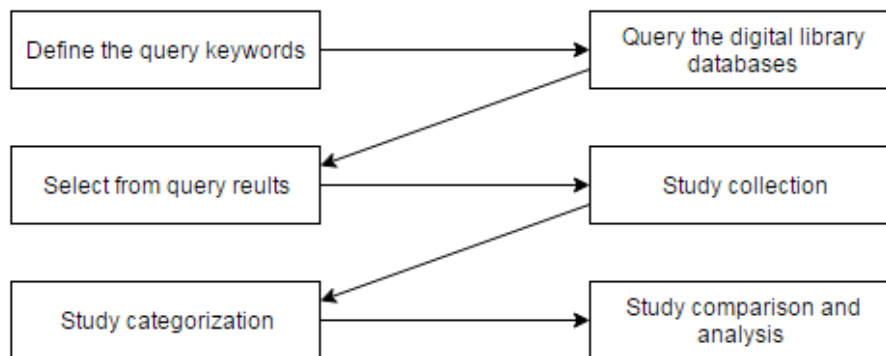
- Requirement 8 (Real-time Reporting) for the *real-time* concept;
- Requirements 4-7 for the *flexible* concept; and
- Requirement 9 (Stakeholder-oriented Views) for the *multi-tenancy* concept.

The next chapter will review existing work related to BI in healthcare, and especially healthcare/pathology dashboards in relation to the above requirements.

## Chapter 4. Literature Review

---

A literature review helps researchers understand the domain, find gaps in existing work, and be evidence-based. The literature review done for this thesis is inspired by the methodology discussed in Kitchenham’s work [28][29]. The process is composed of six steps, shown in Figure 7, and the review for the pathology BI dashboard follows it.



**Figure 7** Literature review process

After a general overview of the usage of BI in healthcare, ten important related academic work and industry solutions are reviewed and analyzed against the twelve requirements identified in the previous chapter.

### 4.1. Overall Usage of BI in Healthcare

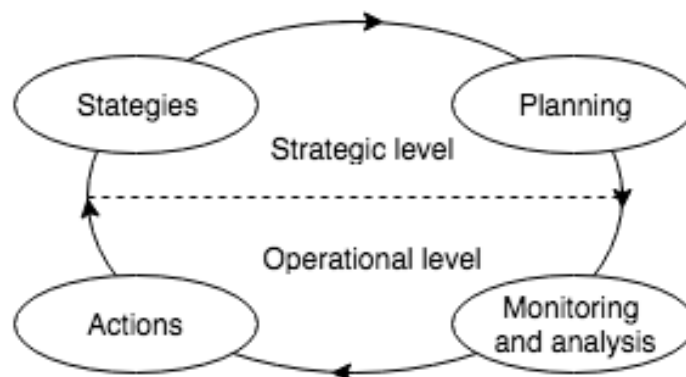
In their report *Crossing the Quality Chasm*, published by the Institute of Medicine (IOM) in 2001 [30], Kohn et al. call to action for six aims to improve the health system in the new century: safe, effective, patient-centered, equitable, timely, and efficient. The last two aims are where BI systems can bring useful contributions. Challenges were raised together with these aims, such as redesigning healthcare processes, and constructing knowledge and skills management. As described in the report, information technology “holds enormous potential for transforming the healthcare delivery system, which today remains relatively untouched by the revolution that has swept nearly every other aspect of society”.

Since then, health information technology (HIT) has been gradually adopted in various areas of health organizations, for example, in the implementation of electronic health records (EHR) [22][43] and health information exchange (HIE) [2][49]. However, making efficient use of the collected and shared data in HIT is still inadequate today. There are many situations where it is very time consuming to search and access data [31]. The application of BI in the healthcare domain comes handy here, with BI's ability to manage, summarize and analyze detailed, historical records [23][45][54].

BI solutions applied in healthcare cover various domains [25], such as:

- Operational, which includes decision support, strategic modeling, workflow management, agility, etc.;
- Financial, which includes enriching revenue cycles, cost management, claim adjudication, etc.;
- Quality, which includes patient satisfaction, customer relationships, quality performance, etc.;
- Public health, with includes disease management, population heal management, epidemiological surveillance, etc.

Using BI in healthcare operations and performance management is one area where there is much research already done [14][35]. By adopting BI solutions in operations management, BI constructs a loop that links strategies, planning, monitoring and analysis, and actions together and merges both the strategic level and the operational level, as shown in Figure 8.



**Figure 8** BI in healthcare operation and performance management loop (adapted from [35])

Applications related to BI in healthcare operations and performance management have been already implemented as well. For example, Nagy et al. [38] describe a web-based BI dashboard that is used in the operational management in a healthcare radiology facility, which helps the department to improve its service performance and productivity.

## 4.2. Keywords, Queries, and Study Collection

For the collection of relevant studies, search keywords were initial focusing on “pathology business intelligence dashboard”, but because of the very limited amount of results returned by the queries, the search domain was expended by including “clinic” and “healthcare” terms. Ovid (including Embase, Cochrane and MEDLINE), IEEE Xplorer, ACM Digital Library, and Scopus have been used as target digital library databases (hence providing coverage of the medical and information systems areas), and the final search string used to query the databases is:

```
( pathology OR pathologist OR healthcare OR "health care" OR
  Clinic )
AND
dashboard
AND
( "business intelligence" OR BI )
```

The selection of articles was done by evaluating each paper’s title and abstract, and only full articles discussing dashboard technical details were kept. Out of several hundred candidates, only six articles [11][13][34][37][39][50] were selected for a comprehensive review. A few popular industry solutions [3][4][6] were also selected to be reviewed, from the top results returned by Google when searching for "pathology dashboard". The work of PowerJ[16] was also selected because it describes the pathology dashboard work previously done at EORLA before the thesis study.

Note that although the work of Chen et al. on real-time dashboards for pathology processes is directly relevant [5], it was not selected because only a short abstract was published, without any technical detail about the solution. The abstract in Chen et al. has confirmed however that the “[c]reation of a comprehensive digital [Quality Dashboard] enabled us to monitor overall laboratory performance in real-time, and served as a guide for quality and process improvement activities”.

### 4.3. Study Categorization, Comparison and Analysis

For the last two steps in the review process shown in Figure 7, the dashboard types and chart types discussed in Section 2.1 are used to categorize the selected studies, while the requirements introduced in Table 5 are used as criteria to evaluate these studies.

#### 4.3.1 Overview and Categorization

A summary of the categorization and analysis results is shown in Table 6.

**Table 6** Summary of the literature review of the pathology BI dashboard.

Article/Product	Dashboard Type	Chart Type	Tabular Data	Data Visualization	Value Filtering	Real-time Reporting	Colour Indicators	Sorting	Drill Through	Interface Customization	Stakeholder-oriented Views	Specimen-type Customization	Automatic Notifications	Workload Planning
McLaughlin et al. [34]	Strategic Tactical	Line chart	Y	Y	N	N	Y	N	N	N	N	N	N	N
Nagy et al. [37]	Tactical Operational	Scatter chart Bubble chart Circular area chart	Y	Y	+/-	N	N	N	Y	N	N	N	N	N
Nash et al. [39]	Tactical Operational	Line chart Column chart Bar chart	Y	Y	N	+/-	N	N	Y	N	N	N	N	N
Wadsworth et al. [50]	Strategic Tactical Operational	Line chart Column chart	Y	Y	Y	N	Y	N	Y	N	N	N	N	N
Ferranti et al. [11]	Tactical Operational	Line chart Column chart Scatter chart	Y	Y	Y	N	+/-	N	Y	N	N	N	N	N
Fox et al. [13]	Operational	N/A	Y	N	Y	N	Y	Y	N	N	N	N	N	N
U. Michigan [6]	Strategic	Line chart Column chart	Y	Y	N	N	N	N	N	N	N	N	N	N
Viewics Inc. [4]	Tactical	Line chart Column chart	Y	Y	N	+/-	Y	N	N	N	N	N	N	N
Kofax [3]	Tactical Operational	Line chart Column chart Bar chart	Y	Y	Y	+/-	Y	Y	Y	Y	N	Y	N	N
PowerJ [16]	Tactical	Line chart Column chart	Y	Y	N	+/-	+/-	N	N	N	N	N	N	N

The first six articles (rows) in the matrix shown in Table 6 refer to dashboards discussed in the literature but not necessarily available on the market.

- McLaughlin et al. [34] introduced a quality dashboard developed for the University of California Los Angeles (UCLA) Department of Neurosurgery. The dashboard gathers patient satisfaction, quality and safety, and operations information, and assists managers with monitoring impact and improving strategies.
- A web-based clinical dashboard described by Nagy et al. [37] is designed to aggregate and display clinical operation data to the surgical managers. The dashboard has enhanced data visualization, and can act as a decision support tool to help management with the continuous quality improvement.
- The reporting tool used by the Ohio State University Medical Center (OSUMC) is discussed in Nash et al. [39]. This tool includes a dashboard, a scorecard, and detailed reports, and provides a data-driven strategy for patient experience improvement.
- A case study of using a business intelligence dashboard at the Cleveland Clinic is described by Wadsworth et al. [50]. This dashboard is designed for the performance management program, and helps improve healthcare quality while reducing cost for the organization.
- A web-based performance dashboard used by the Duke University Health System (DUHS) is discussed in Ferranti et al. [11]. This tool has a dashboard that aggregates information together and also supports drill through to check details. It provides safe access to confidential information and helps users to identify systemic problems.
- In Fox et al. [13], a new panel for displaying healthcare data is discussed. It enables the easy and flexible view of patient data by providing filtering, sorting and colour indicators in the dashboard.

While the above approaches address basic stakeholder requirements such as the need for data to be presented both in a tabular and a graphical manner, they lack a number of key capabilities for pathology management such as the ability to specify different business rules for different types of specimens.

The next three rows in Table 6 refer to commercial pathology dashboards.

- The Clinical Pathology Dashboard for the University of Michigan’s Department of Pathology [6] generates management information that includes pathology operation data, for example, turnaround time for specimens and aggregate monthly data, to provide a high-level overview of the operations during the last month compared to historical data.
- The Anatomic Pathology Solutions, developed by Viewics Inc. [4], is a set of dashboards and reports concerned with quality, productivity, and workload information. By aggregating data of different granularities, this tool brings insight to management about quality and productivity improvement. There is however no clear evidence to support its claimed real-time data capture.
- Another example is the pathology graphical dashboard developed by Kofax [3]. This dashboard supports sorting, filtering, and drilling through, and supports monitoring of the operations and identification of problem areas. However, the refresh of Kofax’s dashboard is manual and hence not truly real-time.

While these commercial tools were specifically developed to address the needs of pathology, none of them addresses all of the requirements identified as key to supporting pathology management, including the ability to: support real-time reporting, modify the interface according to varied stakeholder groups; notify users of upcoming issues; and support workload planning.

The last row in Table 6 refers to a process-level dashboard named *PowerJ*, which is designed to fulfill process and management needs at the Department of Pathology and Laboratory Medicine (DPLM) of The Ottawa Hospital (TOH). This dashboard is mainly concerned with the need for pathology process data to be rendered in a visual manner in order to enable users to quickly grasp the daily performance of the DPLM.

### **4.3.2 Comparison and Analysis**

As can be seen from Table 6, most of the tools used in healthcare BI are tactical dashboards, and some of them support operational dashboards as well. This is because these tools often aim to assist one department of a healthcare organization, for example the pathology department, so the stakeholders of the dashboards will not cover the executive level, unless the tool is used organization-wide, e.g., [6][11][34].

It is again confirmed in these studies that using tables and charts to present data are basic requirements for BI dashboards, and all approaches being discussed in this chapter meet this requirement, except Fox et al. [13] since this is only an operational dashboard which may not have had requirements for data visualization. As for the chart types, line charts, column charts and bar charts are commonly used in these dashboards, because for healthcare organizations, time is an important variable to measure KPIs (line chart) and breaking down KPIs along different categories, e.g. case types, is also a relevant analysis representation (bar charts or column charts). This is in line with what was presented in Table 3. Some other chart types are used in these dashboards, for example, scatter charts and bubble charts.

Colour indicators and value filtering are implemented by half of the related dashboards, however, dynamic sorting is only available in two of them [3][13].

For real-time reporting, three approaches [3][4][39] can support it at some limited level, which means a dashboard can update itself with refreshed data.

Another interesting finding, together with the dashboard type, is about the drill-through capability of dashboards; if a tool supports both tactical and operational views, then drilling through is most likely being enabled. The rationale here is that tactical dashboards normally display aggregated data while operational dashboards are for disaggregated detail data, e.g., event records. When a user is at the tactical dashboard and wants to have some further investigation of detailed operation data, “drilling through” is the right approach to take for the user to achieve that goal. Approaches [3][11][37][39][50] all support drill-through capabilities.

Because of the limitation of the physical area that a dashboard can use (one screen typically), it becomes important for a user to flexibly change the chart type for a dashboard component, and possibly get new insight from different chart types. This interface customization feature is only available in Kofax [3]. Kofax is also the only product that supports specimen-type customization

For the remaining requirements (stakeholder-oriented views, built-in notifications and workload planning), while these three features can make a dashboard more powerful and bring in additional business insight to end users, none of the reported works supports them.

PowerJ [16], the existing BI prototype used at DPLM satisfies the *Tabular Data* requirement by exporting results to a set of MS Excel spreadsheets. The *Data Visualization* requirement is met through the use of line or bar graphs to show the daily states and monthly trends regarding the number of cases, slides, or blocks, and the pathologist workload. *Colour Indicator* is partially fulfilled by using colour to convey the state of cases, blocks and slides in a binary manner (red vs. green for pending vs. ready). *Real-time reporting* is partially fulfilled by updating the dashboard's data in two-hour intervals rather than in much shorter periods.

The PowerJ prototype was evaluated for a period of three months. This trial period showed promising results in terms of turn-around time for processing surgical specimens, but led to the identification of a number of new requirements such as the ability to track the status of specific specimens in real-time in order to identify delays and backlog more rapidly [16]. Real-time here means automatic updates configurable to a refresh rate fast enough to support informed tracking and decision-making (e.g., in the order of minutes or even seconds). In addition, while PowerJ used colour to indicate status, it did not do so following clear business rules such as the time that specimen should be allowed to spend at each stage of the process.

#### **4.4. Chapter Summary**

In this chapter, we first discussed the overall usage of BI in healthcare, and then performed a literature review for healthcare/pathology BI dashboards. The review covers six research articles and three commercial solutions, together with the previous work done for the EORLA pathology BI project (namely, PowerJ). These ten existing approaches were categorized against dashboard and chart types, and evaluated against twelve important criteria. In conclusion, we found that none of the selected research approaches or commercial solutions entirely fulfill the identified requirements, with a large gap in terms of real-time reporting, stakeholder-oriented views, automatic notifications, and workload planning. This research gap motivates the new approach proposed in this thesis, which will be discussed in the following chapters.

## Chapter 5. Architecture

---

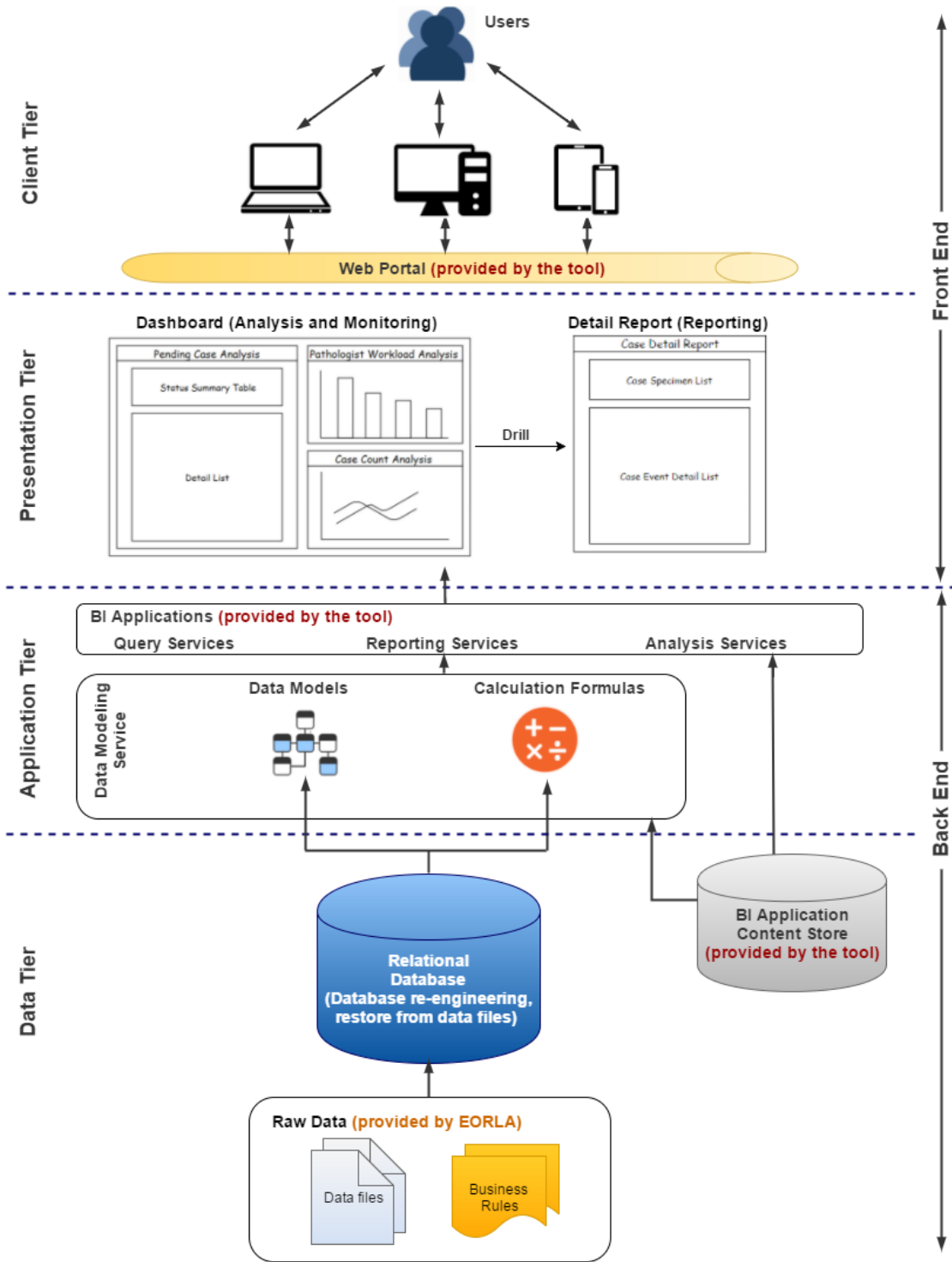
At this point of the DSRM process (Design and Development, see Figure 1), a high-level architecture of our solution is proposed based on the general architecture discussed in Section 2.1.3. This architecture will guide the design and development of the solution with specific technologies (Chapter 6). Although this abstract architecture was developed specifically for the EORLA context, it is generic enough to be reused in other pathology workflow management contexts.

### 5.1. Abstract Architecture of the Solution

With the emergence of dashboards as tools for integrating and monitoring KPIs, a real-time monitoring technique can automatically provide relevant information in a timely fashion that could improve the quality and efficiency of the pathology work process management. Real-time access to information and indicators enables optimizing the process from case accession to a diagnostic report and alert management when a link of the process chain is bottlenecked.

Following this business strategy, a high-level abstract architecture of the solution is designed to guide solution development and the implementation process. This architecture is built based on the general BI architecture discussed in Section 2.1.3, with some component added or eliminated specifically for this project. Figure 9 shows the EORLA BI project abstract architecture diagram contributed by this thesis.

The rest of this chapter will describe the content of its four tiers, from bottom to top.



**Figure 9** EORLA BI project abstract architect diagram<sup>5</sup>

<sup>5</sup> Highlighted component in the diagram, such as provided by the tool or provided by EORLA, means the component is already available to the study and thus do not need design and implementation.

## 5.2. Data Tier

Compared to the general BI architecture (Figure 2), the major change in the *data tier* of the EORLA BI project is an additional component to process the raw data.

In a general BI architecture, the data source database is usually set up together with the external ETL (extract, transform, and load) set up, and data stored in the database is handled by the ETL process. For example, in EORLA's production area, the local clinical database (MySQL©) is setup together with the barcode scanning system (PowerPath®), and the operational data of the pathology work process is captured and stored back to the database by each barcode tracking scan.

Due to the fact that the confidential information (patient information) stored in the EORLA clinic database needs to be protected, the EORLA BI project does not (yet) have direct access to the production database. Instead, one month of the department operation data (without confidential information) was exported to comma-separated value (CSV) files. In addition, business rules were provided in Excel spreadsheet pivot tables. In order to make the raw data understandable and accessible by the BI applications in this kind of situation, database re-engineering is required to restore a database from the data files, and the related development work includes:

- Analyzing the given data files, and then developing the “create” database and create table statement to restore the database;
- Data cleaning in the given data files, and then loading the data files to the re-stored database.

Because the restored database is built from limited data files, additional tables might be added in order to fulfill technical requirements while working on the implementation of other tiers.

Note that such situation is common when researchers are involved in the development of healthcare systems. Once the proof-of-concept is satisfactory, the production system can be used directly as a data source instead of this file-based infrastructure.

## 5.3. Application Tier

### Service Selection in BI Application Services

The decision of what BI application services to select for the EORLA BI project was made based on how best to fulfill the user requirements while considering priorities and the resources available (including research time). In this context, a query service, a reporting service, and an analysis service were selected to build a fully functional BI solution. Compared to the BI applications from the general architecture, one will find that predictive analysis is not included in our project. Accordingly, Requirement 12 from Table 5 (Workload Planning) becomes out of scope for this thesis's tool. The major reason behind this decision is that, although predictive analysis can add valuable insights to the solution and help the pathology management determine where the future performance bottleneck could be located, predictive analysis is usually not available out-of-the-box in many BI tools and also requires much configuration.

### OLTP vs. OLAP Selection in Data Modeling

One will also find that the OLAP cube component is removed in the EORLA BI architecture *Application Tier*. Although it is commonly suggested to adopt OLAP in BI dashboard solutions (see Section 2.1.6), OLTP was selected for this project because of the following reasons:

- The pathology department's operational data is stored in an OLTP database only, and there is no data warehouse set up for storing data in a multi-dimensional data structure. Therefore, the restored database is set up in the same way, to simplify future connections to the operational database. To use OLAP cubes in this case, EORLA could set up a data warehouse in the data tier, but this option is not considered for this study because it is too costly in terms of product license fees and development/maintenance effort. An alternative is to build the cubes from the relational database at data modeling time, but this could result in complicated development work when building cubes and poor performance when running cubes.
- Real-time monitoring requires high-speed data retrieval against the ongoing operational data. Querying directly the transactional tables offers better performance than OLAP cubes, which takes longer time on fitting data to the multi-

dimensional data structure and also adds extra burden to the operational database.

- In order to quickly query case details, OLTP is designed for this requirement while OLAP is more suitable for summarized, aggregated, crosstab, or chart analysis.
- One technical detail requirement we got from the EORLA is to filter out weekends and be focused on the current month only for the case count analysis. Both OLTP and OLAP can achieve this requirement. However, using OLAP does not bring benefits in this case. Organizations usually want to analyze their business along different time intervals, such as daily, monthly, quarterly, and yearly, and OLAP provides simple support for this by building a hierarchical date dimension and allowing users to drill up and down. Because drilling up/down along the time dimension is not a requirement of the EORLA BI project, OLAP support is not really required.
- Dynamic filtering is one of the project requirements, and both OLTP and OLAP support it. However, the performance of OLAP can be very poor when there are multiple dynamic filters required. The reason behind this issue is that tables are de-normalized in OLAP and filters showing in the dashboard are usually applied on the content field not the keys. To support the filters, the fact table has to join the dimension tables and this can be very expensive and bring poor performance.

## 5.4. Presentation Tier

After discussing with the end users, it was decided that the *presentation tier* would have two components. The first one is a BI dashboard, which will support real-time monitoring on pending cases, and will analyze pathology workloads and department case counts. The second component is a BI report, which is drill-linked to the dashboard and presents the detailed information of a case, such as the specimens within that case and the case event information.

## 5.5. Client Tier

As shown in Figure 9, there is not too much development work required in the *Client Tier*. Unlike the PowerJ project, which is a desktop-based application programmed in Java, the solution developed in this thesis is web-based, which has advantages over a desktop-based solution:

- A web-based approach is a lightweight thin-client solution, which means all the BI software is centralized in only one or a few servers (in the case where the BI software and the database software are deployed on separate nodes). The end users just need to have a web server and internet access (or an intranet, depending on whether the BI web portal is opened publicly or is privately accessible) access on their devices.
- The solution is easily accessible by multiple users and supports different devices, such as desktop, laptops, or even mobile devices. This is in line with the flexibility concept expressed in the research problem (Section 1.2).

## 5.6. Chapter Summary

This chapter provided an abstract architectural view of the solution developed in this thesis, together with the rationale for several architectural decisions taken in each of the architecture's four tiers. This architecture guides the design and implementation of the solution, which will be discussed in the next chapter.

## Chapter 6. Development and Demonstration

---

The development and demonstration stages of the DSRM process are closely interleaved in the EORLA pathology BI project. The reason is that, in a BI solution development, developers need to present prototypes in certain contexts to verify if the solution can answer business questions. In this research, because historical data was available in data files, we had the solution directly running on this data set, which led to useful ongoing feedback along multiple prototype versions. The content in this chapter is organized according to the tiers defined in the abstract architecture from the previous chapter, but starts with a brief discussion of the technologies selected for the implementation and ends with additional details related to the real-time nature of the BI solution.

### 6.1. Tools Selection

In order to implement the abstract architecture and build a BI solution for the EORLA BI project, tools and technologies need to be selected for each tier. The selection here is based mainly on the presentation/application tier, where BI tools usually reside. As shown in Table 7, IBM Cognos 10 was chosen here mainly because of its flexibility at the user interface (UI) level, its large market share, and the familiarity of the thesis author with this tool. IBM Cognos works very well with IBM DB2 in the data tier, and Microsoft Internet Explorer for the client tier. The latest versions available at the time were used. These tools are not the only ones available, and Table 7 shows alternatives that could have been used.

**Table 7** List of selected tools

	<b>Selected Tool</b>	<b>Alternative Tool</b>
<b>Client Tier</b>	Internet Explorer 11	Firefox, Chrome
<b>Presentation Tier Application Tier</b>	IBM Cognos 10 Business Intelligence Suite	Tableau Desktop, Oracle BI
<b>Data Tier</b>	IBM DB2 10.5	MySQL, Oracle Database

## 6.2. Data Tier Development

### 6.2.1 Restored Database

As discussed in Section 5.2, the researcher did not have direct access to the production database and needed to restore a database from data files. This new database does not need to reflect the structure of the operations database. One month (January 2015) of the department's operation data (without confidential information) were exported to CSV files and allowed to be used in the project. Characteristics of these sample data files can be found in Table 8.

**Table 8** Sample data files characteristics

<b>File Name</b>	<b>Column Count</b>	<b>Row Count (Header row included)</b>
acc_amp_event.csv	13	289,316
acc_block.csv	14	26,451
acc_order.csv	24	108,380
acc_process_step.csv	5	31,712
acc_specimen.csv	31	10,961
acc_type.csv	16	27
accession_2.csv	45	6,549
lab_procedure.csv	17	923
process_step.csv	11	334
tmpl_profile.csv	10	1361

Except for the data files, there was no database schema or table information provided, for example, the CREATE DATABASE or CREATE TABLE database statements. As the data files included field names in the first (header) row, a sample database was restored locally, following the rules listed below in the defined order.

- If a column name contains “\_id”, define the column data type to be INTEGER;
- If a column contains only (1) “Y” or “N” (case insensitive), or (2) 1 or 0, define the column data type to be CHAR(1);

- If the content in a column is in date or time format, define the column data type to be `TIMESTAMP`;
- If a column contains digital only, define the column data type to be `INTEGER` if there is no decimal point, otherwise set it to `DOUBLE`;
- For all the rest columns, define the data type to be `VARCHAR(255)`.

No hypothesis was made on the primary key, foreign key, or table constraints of the physical tables, because building a data model on physical tables with assumptions on joint relationships may cause serious conflicts and failure when deploying to a production database that has no joint or different joint relationships. Table relationships were defined during data modeling at the application tier, and the related design decisions and implementation can be found in Section 6.3.

Some manual data cleaning was done before loading sample data files to the local database, for example:

- Replace “NULL” by empty content, because “NULL” is a string saved in a CSV cell, and will cause a data type error if loaded to a non-string field;
- Re-format date or time columns to a unified format “YYYY-MM-DD HH:MM:SS”, as this is the default timestamp format that the IBM DB2 load utility can recognize<sup>6</sup>. By setting the date or time to the desired format, there is no need to add `timestampformat="x"` in each table definition if it contains a timestamp column, which will help simplify the `CREATE TABLE` statement and lower the risks of incorrect data format during data load.
- Remove the header row in each data file. Unlike DB2’s `IMPORT` command, which can skip headers in the data files by using `SKIPCOUNT`, this feature is deprecated in the `LOAD` command. However, `LOAD` was still selected in this case study because it has a better performance compared to `IMPORT`<sup>7</sup>.

---

<sup>6</sup> DB2 10.5 for Linux, UNIX, and Windows, `LOAD` command, [https://www.ibm.com/support/knowledgecenter/SSEPGG\\_10.5.0/com.ibm.db2.luw.admin.cmd.doc/doc/r0008305.html](https://www.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.admin.cmd.doc/doc/r0008305.html)

<sup>7</sup> DB2 10.5 for Linux, UNIX, and Windows, Differences between the import and load utility, [https://www.ibm.com/support/knowledgecenter/SSEPGG\\_10.5.0/com.ibm.db2.luw.admin.dm.doc/doc/r0004639.html](https://www.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.admin.dm.doc/doc/r0004639.html)

## 6.2.2 Additional Data Tables

A few additional tables were added to the restored database in addition to the tables created from data files. The first one is for business rules. The pathology operation process business rules were provided in an Excel spreadsheet as pivot tables, which have multiple heading rows and cannot be directly loaded to a database table. Data was cleaned and saved in a CSV file in this case, and a table was created accordingly.

As discussed in Section 5.2, extra tables might be added through the development of tiers other than the data tier, and this requires to work back to the data tier. The second newly added table, the pathologists table, was such a case. The demand for this table came from the implementation of the presentation tier, which requires showing the pathologist names or name initials in a chart's x-axis. Though the production database has the pathologist table, no data was provided by the hospital due to the confidentiality of this information. To resolve this problem, one single table with only two columns was created in the sample database. While the data of the pathologist ID field came from "select distinct assigned\_to\_id from acc\_process\_step", anonymous names "P1, P2, ..., Pn" were assigned to each ID manually in order to have distinct and confidential identifiers.

The CREATE TABLE statements formalizing the content discussed in Sections 6.2.1 and 6.2.2 can be found in Appendix G.

## 6.3. Application Tier Development

### 6.3.1 Data Model Implementation

The data modeling implementation in this tier has two layers – the physical layer and the logical layer. In the *physical layer*, tables in the underlying database (the restored database here) were directly added to the (Cognos) data model by using SQL (Structured Query Language) queries such as "select \* from table\_name". Join relationships and cardinalities are imported to the physical layer if there exists any in the database, and this is why no such items were added in the restored database, as discussed in Section 6.2.1.

The *logical layer*, also called the business layer, is built upon the physical layer. Two tasks were performed in this layer. The first one was to assign human readable names

to each table in order for them to become user-friendly, and the second one was to determine the usage of each table and add required table relationships for the BI dashboard.

A summary of the data model implantation work can be found in Table 9 and the full logical layer diagram is included in Appendix H. Note that this data model is internal to the BI tool and is used to perform indirect queries on the underlying (restored) database.

**Table 9** Data model implantation work summary

Physical Table Name	Logical Table Name	Usage	Join Relationships
acc_amp_event	Accession Event	Stores events where scanning barcode for case/specimen/block /slide that triggers an event	[Accession Event].[acc_id] = [Accession].[id], n to 1
acc_block	Block	Stores data about each block of a specimen	[Block].[acc_specimen_id] = [Specimen].[id], n to 1
acc_order	Slide	Stores data about each slide of a specimen	1. [Slide].[acc_specimen_id] = [Specimen].[id], n to 1 2. [Slide].[procedure_id] = [Slide Type].[id], n to 1
acc_process_step	Accession Process	Stores each process record of a case	1. [Accession Process].[acc_id] = [Accession].[id], n to 1 2. [Accession Process].[step_id] = [Process Step].[id], n to 1 3. [Accession Process].[assigned_to_id] = [Pathologist].[id], n to 1
acc_specimen	Specimen	Stores data about each specimen of a case	1. [Specimen].[id] = [Slide].[acc_specimen_id], 1 to n 2. [Specimen].[id] = [Block].[acc_specimen_id], 1 to n 3. [Specimen].[acc_id] = [Accession].[id], n to 1 4. [Specimen].[tmplt_profile_id] = [Specimen Type].[id], n to 1
acc_type	Case Type	A user-defined dictionary that stores case types	[Case Type].[id] = [Accession].[acc_type_id], 1 to n

Physical Table Name	Logical Table Name	Usage	Join Relationships
accession_2	Accession	Stores data about each case (one entry per case) that is sent to a pathologist for a diagnosis	1. [Accession].[id] = [Specimen].[acc_id], 1 to n 2. [Accession].[id] = [Accession Process].[acc_id], 1 to n 3. [Accession].[id] = [Accession Event].[acc_id], 1 to n 4. [Accession].[acc_type_id] = [Case Type].[id], n to 1
lab_procedure	Slide Type	A user-defined dictionary that stores slide types	[Slide Type].[id] = [Slide].[procedure_id], 1 to n
process_step	Process Step	A user-defined dictionary that stores operational processes in the pathology department	[Process Step].[id] = [Accession Process].[step_id], 1 to n
tmplt_profile	Specimen Type	A user-defined dictionary that stores specimen types	1. [Specimen Type].[id] = [Specimen].[tmplt_profile_id], 1 to n 2. [Specimen Type].[type] = [Business Rules].[specimen_type], 1 to 1
business_rules	Business Rules	A user-defined dictionary that stores business rules	[Business Rules].[specimen_type] = [Specimen Type].[type], 1 to 1
pathologist	Pathologist	A look-up table that stores pathologist IDs and anonymous names	[Pathologist].[id] = [Accession Process].[assigned_to_id], 1 to n

### 6.3.2 Calculation Formulas

Mathematical formulas were built in the data model, based on business rules and the information provided by the pathology department's operation manager. The list below reflects the final set of calculation formulas used in the EROLA BI prototype. In general, after receiving business rules, a prototype of the calculation formulas was generated to validate and identify additional gaps in the rules. New prototypes were developed by fixing/adding business rules, until the set of rules became stable and approved by managers.

1. Target time of a specimen type:

$$T_{specimen\_type\_target\_time} = \sum_{i=1}^5 [Business\ rules]. [stage_i],$$

where *stage1* = accession time to gross time,

*stage2* = gross time to embeded time,

*stage3* = embeded time to microtomy time,

*stage4* = microtomy time to routing time,

*stage5* = routing time to signout time

2. Target time of a case:

$$Target\ h. = \max_u T_{specimen\_type\_target\_time}(u),$$

for specimen\_type *u* in a case

3. Spent time of a case:

$$Spent\ h. = \Delta time = T_{current} - T_{case\_open},$$

where  $T_{case\_open} = [Accession]. [created\_date]$

4. Ratio:

$$ratio = Spent\ h. / Target\ h. \times 100\%$$

5. Case colour code (the 70% value was agreed to by the stakeholders):

$$colour\ code = \begin{cases} red, & \text{if } ratio \geq 100\% \\ amber, & \text{if } 70\% \leq ratio < 100\% \\ green, & \text{if } ratio < 70\% \end{cases}$$

### 6.3.3 Aggregation, Sorting, and Parameters

Aggregation, sorting, and parameters required by the presentation tier development were implemented in the application tier, and details are shown in Table 10.

**Table 10** Aggregation, sorting and parameters in the dashboard components

<p><b>Pending Case Analysis Component</b></p> <p><b>Aggregation:</b> Total pending case count by lab and by colour code; Total pending case count overall by colour code</p> <p><b>Sorting:</b> Sorted by the case created date ascending</p> <p><b>Parameter:</b> p_show, embedded in drop-down list value prompt, filter on colour code; p_lab, embedded in drop-down list value prompt, filter on lab; p_case_type, embedded in drop-down list value prompt, filter on case type.</p>
<p><b>Pathologist Workload Analysis Component</b></p> <p><b>Aggregation:</b> Total pending slide count by pathologist and by case category Case category by time create statement:</p> <pre>case   when ([Time Range]=24 and [Case Type]='Autopsy')     then ('Autopsy, &lt; 24 H')   when ([Time Range]=30 and [Case Type]='Autopsy')     then ('Autopsy, 24~48 H')   when ([Time Range]=48 and [Case Type]='Autopsy')     then ('Autopsy, &gt;=48 H')   when ([Time Range]=24 and [Case Type]='Cytology')     then ('Cytology, &lt; 24 H')   when ([Time Range]=30 and [Case Type]='Cytology')     then ('Cytology, 24~48 H')   when ([Time Range]=48 and [Case Type]='Cytology')     then ('Cytology, &gt;=48 H')   when ([Time Range]=24 and [Case Type]='Surgical')     then ('Surgical, &lt; 24 H')   when ([Time Range]=30 and [Case Type]='Surgical')     then ('Surgical, 24~48 H')   when ([Time Range]=48 and [Case Type]='Surgical')     then ('Surgical, &gt;=48 H') end</pre> <p><b>Sorting:</b> Sorted by the total pending slide count of each pathologist descending</p> <p><b>Parameter:</b> None</p>

<p><b>Case Count Analysis Component</b></p> <p><b>Aggregation:</b>  Total created case count by day;  Total finalized case count by day;  Backlog case count by day, calculation<sup>8</sup>:  <code>running-total([Created Case Count]-[Finalized Case Count])</code></p> <p><b>Sorting:</b>  Sorted by the total pending case count of each pathologist descending</p> <p><b>Parameter:</b>  None</p>
<p><b>Case Detail Report</b></p> <p><b>Aggregation:</b>  None</p> <p><b>Sorting:</b>  Sorted by case event timestamp descending</p> <p><b>Parameter:</b>  p_case_id, embedded in rerport drill-through definition and the select &amp; search value prompt, filter on case accession number</p>

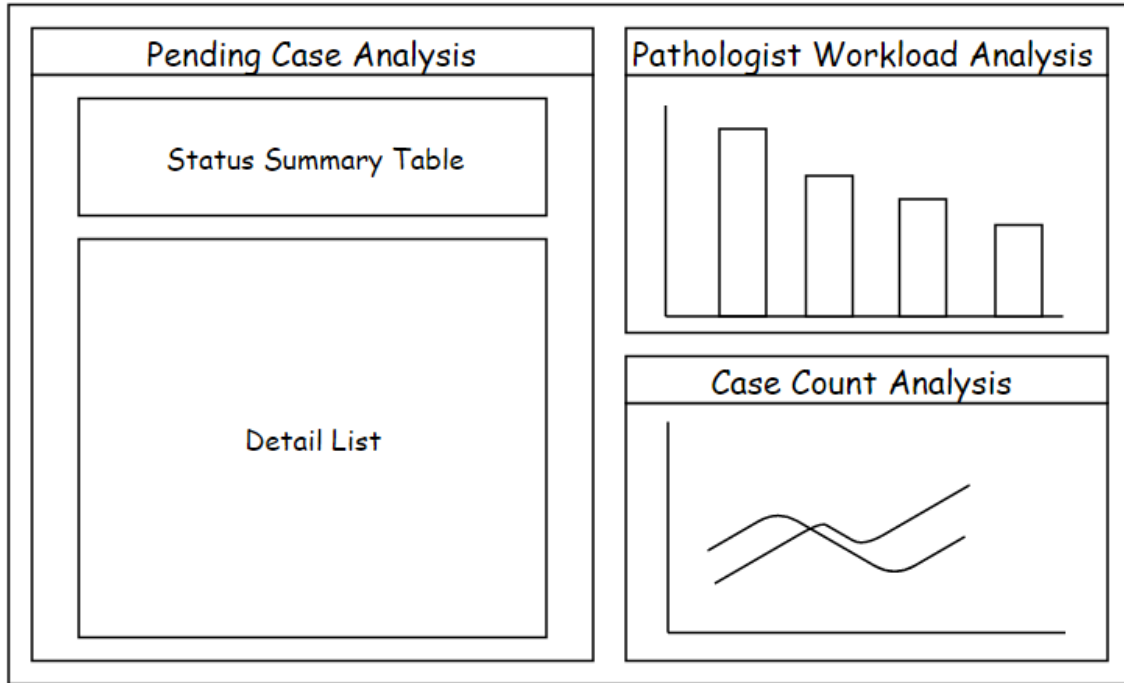
## 6.4. Presentation Tier Implementation

### 6.4.1 Dashboard Layout Design

During the design phase with the end users, it was decided that the dashboard will contain three components, which are the pending case analysis, pathologist workload analysis, and case count analysis. Based on the common reading patterns (left to right) and taking into account the component importance, these components were placed into the dashboard layout draft as illustrated in Figure 10.

---

<sup>8</sup> Cognos Business Intelligence 10.2.2, Author and Create Report Studio User Guide 10.2.2, Using the expression editor, Summaries, running-total. [https://www.ibm.com/support/knowledge-center/en/SSEP7J\\_10.2.2/com.ibm.swg.ba.cognos.ug\\_cr\\_rptstd.10.2.2.doc/c\\_ces\\_runningtotal.html](https://www.ibm.com/support/knowledge-center/en/SSEP7J_10.2.2/com.ibm.swg.ba.cognos.ug_cr_rptstd.10.2.2.doc/c_ces_runningtotal.html)



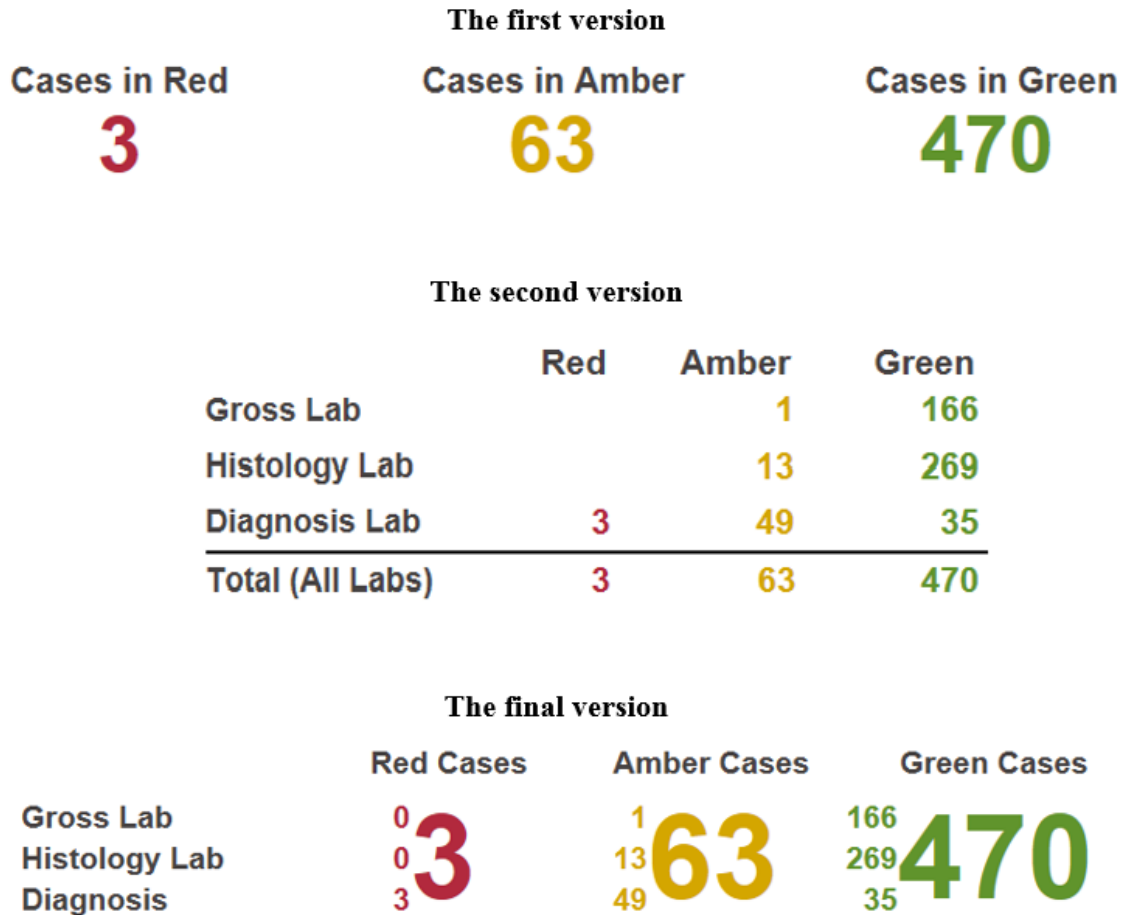
**Figure 10** Dashboard layout

### 6.4.2 Pending Case Analysis Component

The first component in the dashboard is the Pending Case Analysis, which provides real-time monitoring (auto-refreshed every 2 minutes) of pending cases. It helps management answer questions such as “what is the total number of pending cases?”; “at which stage of the process are cases of interest?”; and, “which cases need to be prioritized to avoid exceeding target processing times?”.

#### The Static Section

The upper section of the component is designed to provide the summary information about the status of pending cases, based on target processing times. It acts as an alert system for Operations Manager or Histology Laboratory supervisors, showing the total numbers of pending cases broken down by process stage and laboratory, and enabling a quick understanding of their overall status.



**Figure 11** Pending Case Analysis upper section design

Using colour codes similar to a “traffic light” system is one of the initial requirements from the users, and this was the start point of the design of the upper section as well. Figure 11 shows the design process of this part, composed of three iterations.

In the first version, the component displayed the total pending case count, which is defined by the formula discussed in Section 6.3.2, in different statuses using different case colours.

When meeting with the end users for the design review, they suggested that it would be more meaningful to add the case count for each lab to this section. By doing so, the component will not only be useful to the department’s operation manager, but it can also help the lab managers understand what is the current status in his or her lab. Thus, the

second version was created, which was using a conventional crosstab to display the pending case status in different labs, with a total summary row at the bottom to show the overall status of the whole department.

The final version was an optimized second version. It displays the exact same information as in the second version, but in a better presentation format according to William's contrast, repetition, alignment and proximity (CRAP) design principles [52]. By using different font sizes to differentiate data at a department level versus at a lab level and by eliminating the summary row, the new view became more visually efficient and compact, and saved space to display more information in the lower section.

### **The Dynamic Section**

The lower section of the component, shown in Figure 12, provides a dynamic, detailed list of pending cases that can be filtered by the value prompts. Three drop-down lists on colour code, laboratory, and case type control the details of the list. Each list prompt has four options (as shown in Figure 13), and this prompt group provides a combination of  $4^3 = 64$  different choices to the users for them to quickly find and focus on the cases of interest. There are seven columns included in the list, which were provided by the end users during the design phase as these are the most important fields for the users to check the status and stage of a case at a glance. The list is by default set to be sorted by the case creation date, which is the natural way to view transactions from earliest to the latest. However, the list can be dynamically sorted by any other columns, for example by processing stage (need embedding, need microtomy, need quality control and distribution, need diagnosis, need finalized), or the ratio column. This allows each user group at the pathology facility to personalize the component's display according to their interests and responsibilities.

All Color Codes		All Labs		All Case Types			
Accession No.	Case Type	Process Stage	Created Date	Spent h.	Target h.	Ratio (S h./T h.)	
<a href="#">CY-15-001259</a>	Cytology	Need Diagnosis	15-01-29 08:35 AM	50	64	78.12%	
<a href="#">CY-15-001258</a>	Cytology	Need Diagnosis	15-01-29 08:34 AM	50	64	78.12%	
<a href="#">CY-15-001260</a>	Cytology	Need Diagnosis	15-01-29 08:37 AM	50	64	78.12%	
<a href="#">SP-15-003900</a>	Surgical	Need Microtomy	15-01-29 08:38 AM	50	79	63.29%	
<a href="#">CY-15-001261</a>	Cytology	Need Diagnosis	15-01-29 08:38 AM	50	64	78.12%	
<a href="#">SP-15-003901</a>	Surgical	Need Embedding	15-01-29 08:40 AM	50	178	28.09%	
<a href="#">SP-15-003903</a>	Surgical	Need Embedding	15-01-29 08:48 AM	50	84	59.52%	
<a href="#">SP-15-003902</a>	Surgical	Need Embedding	15-01-29 08:48 AM	50	114	43.86%	
<a href="#">SP-15-003905</a>	Surgical	Need Diagnosis	15-01-29 08:50 AM	50	40	125.00%	
<a href="#">SP-15-003907</a>	Surgical	Need Embedding	15-01-29 08:50 AM	50	130	38.46%	
<a href="#">SP-15-003908</a>	Surgical	Need Embedding	15-01-29 08:52 AM	50	130	38.46%	
<a href="#">SP-15-003909</a>	Surgical	Need Microtomy	15-01-29 08:53 AM	50	79	63.29%	

Top Page up Page down Bottom  
**Legend:** over processing time target – red; within 30% band below target = amber; all others = green  
 Reporting Time: Feb 2, 2015 10:00:00 AM

**Figure 12** Pending Case Analysis lower section overview

Color Code	Lab	Case Type
All Color Codes	All Labs	All Case Types
Red	Gross Lab	Autopsy
Amber	Histology Lab	Cytology
Green	Diagnosis Lab	Surgical

**Figure 13** Pending Case Analysis value prompt options for drop-down lists

For example, the first entry in the lower section of Figure 12 shows a cytology case CY-15-001529 that needs diagnosis. The highlighted amber background in the ratio column indicates that at least one of its specimens has exceeded 70% of its target processing time. In order to access more detailed information about this case, the user can click its Accession Number, and a detailed report then opens in a new window, which will be discussed in Section 6.4.5.

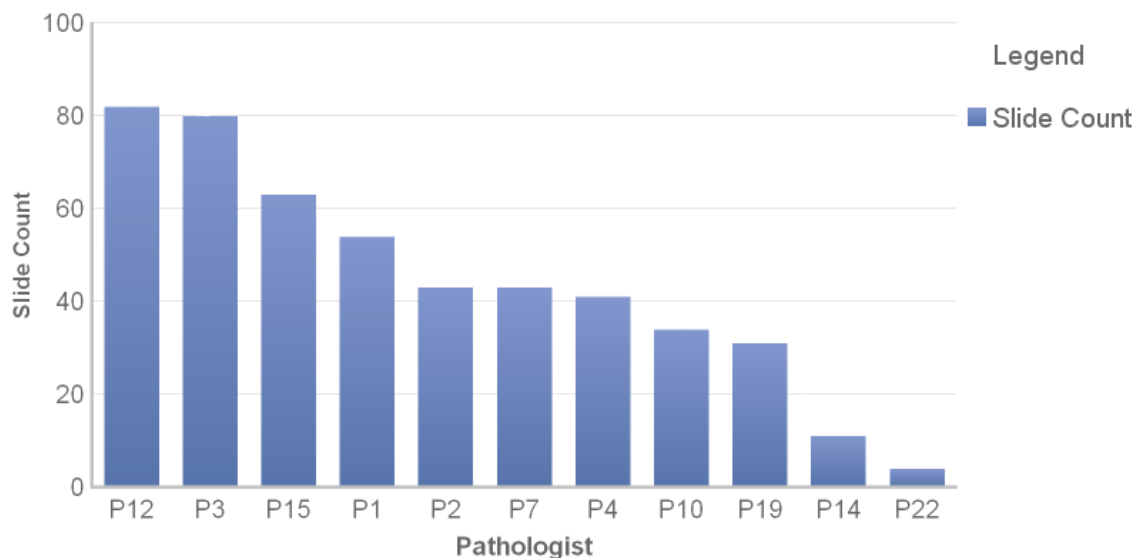
Some additional information can be found at the bottom of the list, as shown in Figure 12. The first piece is the legend of the pending case analysis component, which explains what each colour code means in natural language and is more understandable to the end users than showing the mathematical formulas built in Section 6.3.2. The second piece displays the reporting time. In a real productive environment, it will show the timestamp of the last run or refresh, by using some function provided by the BI tool or the

database vendor, for example `current_timestamp`<sup>9</sup>. In the case study, because the sample data was for January 2015 only, the `[current time]` attribute was hard-coded to “2015-02-02 10:00:00.000”, otherwise the calculated spent hours (Spent h. column) would be extremely large and all the pending cases would be in red if the actual current timestamp was used.

### 6.4.3 Pathologist Workload Analysis Component

The dashboard’s second component (top-right part of Figure 10) is the Pathologist Workload Analysis. This component enables the Chief of the Division and the operation manager to answer questions such as “What is each pathologist’s workload?” and, “Which types of cases is each pathologist currently processing?”

In order to show the comparison of the pending cases count for each pathologist, the bar chart format is selected for data visualization with pre-sort enabled on case count from highest to lowest, as shown in Figure 14.












**Figure 14** Pathology Workload Analysis first version

<sup>9</sup> Cognos Business Intelligence 10.2.2, Model and Design Framework Manager User Guide 10.2.2, Using the expression editor, Common Functions, `current_timestamp`  
[http://www.ibm.com/support/knowledge-center/en/SSEP7J\\_10.2.2/com.ibm.swg.ba.cognos.ug\\_fm.10.2.2.doc/c\\_sql99\\_current\\_timestamp.html](http://www.ibm.com/support/knowledge-center/en/SSEP7J_10.2.2/com.ibm.swg.ba.cognos.ug_fm.10.2.2.doc/c_sql99_current_timestamp.html)

When reviewing this design with the end users, they brought in a new request. They wanted to have a more detailed view of the case type, and preferred to see how long the work has been taking in a summary view. In order to achieve this, a case category by item was created at the backend (discussed in Table 10), and the heat map concept was used in the frontend presentation.

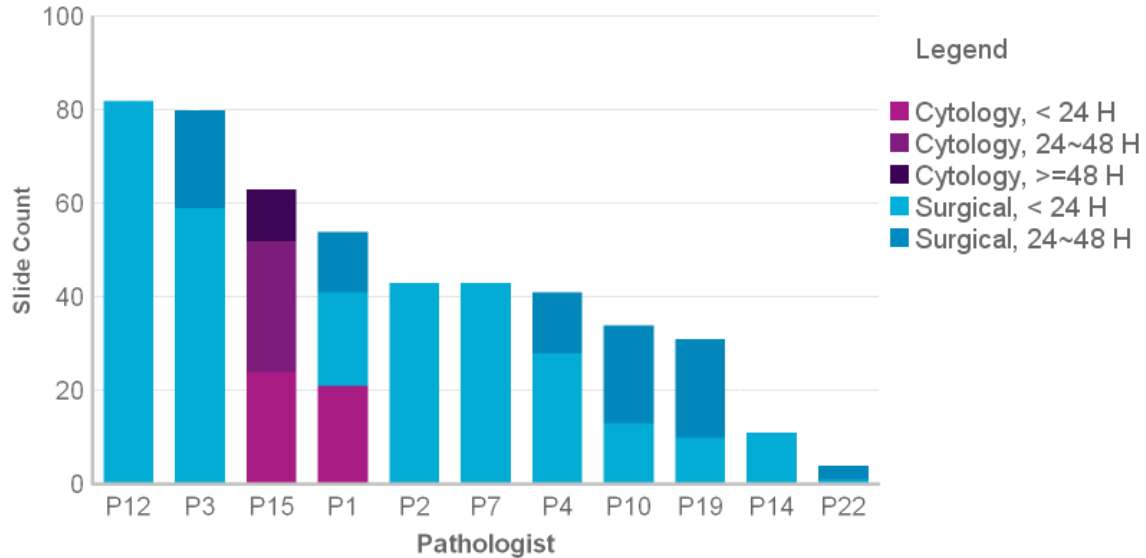
A heat map is “a graphical representation of data where the individual values contained in a matrix are represented as colours” [17]. A customized heat map colour palette for the pathology workload was designed. As illustrated in Table 11, one colour type is assigned to a case type, and within a case type, colours in different tones are assigned to different categories by time based on how long the diagnosis takes.

**Table 11** “Heat map-like” colour palette for Pathology Workload Analysis

Case Type	Colour Type	Case Category by Time	Colour
Autopsy	Yellow	Autopsy, < 24 H	#F5BB00 
		Autopsy, 24~48 H	#EC9F05 
		Autopsy, >=48 H	#BF3100 
Cytology	Purple	Cytology, < 24 H	#AB1A86 
		Cytology, 24~48 H	#7F1C7D 
		Cytology, >=48 H	#3B0256 
Surgical	Blue	Surgical, < 24 H	#00B0DA 
		Surgical, 24~48 H	#008ABF 
		Surgical, >=48 H	#003F69 

After using the case category by time to break down the bar chart, new stacked bar chart was created, as shown in Figure 15 (the legend in the chart is created dynamically by show-

ing only the categories actually used in the diagram). This chart provides abundant information to the users in colour, stack proportion and bar height, and users can easily get answers to their related business questions by analyzing this component.



**Figure 15** Pathology Workload Analysis final version

#### 6.4.4 Case Count Analysis Component

The third component in the dashboard (bottom-right part of Figure 10) is the Case Count Analysis (Figure 16). This component was inspired by the work done in the PowerJ prototype [16], and also includes more information about the backlog case count as requested by the users. It summarizes two distinct types of information: the number of backlog cases, and the amount of created vs. finalized cases. Since the scale for the case count differs from that of the backlog case count, there are two y-axes set in the chart. The y-axis on the left, in blue, is the primary axis and is used for the backlog case count, while the secondary axis on the right, in grey, is used for created and finalized case counts. This component thus helps the operations manager and the clinical management of DPLM to answer the following questions: “How many cases are created and finalized daily over a month?” and “How many cases have been backlogged in the past month, and what is the trend?” This layering of information allows users to understand if, for example, a widening gap between created

and finalized cases is due to a sudden increase in incoming cases, or due to a gradual accumulation of cases to be processed because of the demand being beyond current capacities.

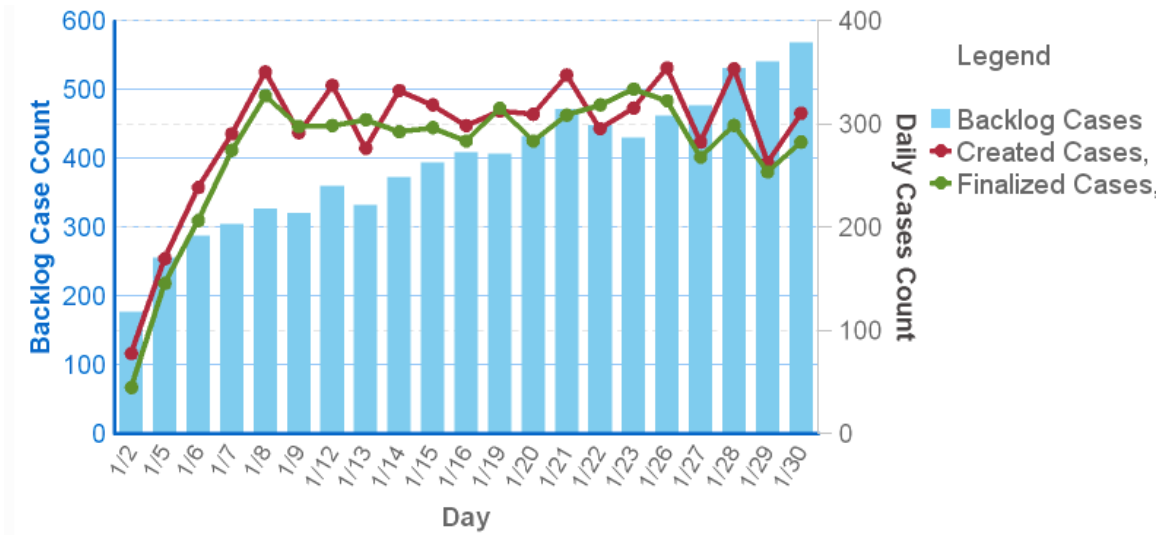


Figure 16 Case Count Analysis

### 6.4.5 Case Detail Report

In addition to all the components contained in the dashboard, a case detail report is also included in the solution as a drill-through target report of the pending case analysis. When a user clicks an Accession Number in the pending case analysis list (Figure 12), the case detail report opens in a new window and displays all the specimens and time-stamped events related to it, as shown in Figure 17. Events are pre-sorted from latest to earliest, so users can quickly get what is the latest status of a case and where to find this case if any action is required.

Accession No.: SP-15-003905

Specimen ID	code	description	specimen_label	recv_date	billable	service_code_id	created_date	last_final_collection_date	Target H.
1562543	KIDT	right transplant kidney	1	29/01/15 8:49 AM	N	75	29/01/15 8:50 AM	29/01/15 12:00 AM	40

event_date	amp_mode	event_type	event_description
30-Jan-15 1:07:02 PM	Diagnostician	folder_scanned	Received in folder by Blanco MD, Paula.
30-Jan-15 1:07:02 PM	Diagnostician	folder_scanned	Received in folder by Blanco MD, Paula.
30-Jan-15 1:07:02 PM	Diagnostician	folder_scanned	Received in folder by Blanco MD, Paula.
30-Jan-15 1:07:02 PM	Diagnostician	folder_scanned	Received in folder by Blanco MD, Paula.
30-Jan-15 1:06:55 PM	Diagnostician	material_scan	Scanned by Blanco MD, Paula.
30-Jan-15 12:26:37 PM	Slide Distribution	material_routed	Routed to Blanco MD, Paula by Keyes, Lauren.
30-Jan-15 12:26:37 PM	Slide Distribution	material_routed	Routed to Blanco MD, Paula by Keyes, Lauren.

Figure 17 The case detail report

The detail report can also work as a stand-alone report for rapid case-querying checks. For example, if a user wants to know what is the current status of case SP-15-003905, he or she can directly open this case detail report and type in the Accession No. in the value prompt text box (even just part of the Accession No. will work as shown in Figure 18). After selecting the desired case number and clicking the Finish button, the case detail report is displayed to the user as well.

**Accession No.**

Keywords:  
Type one or more keywords separated by spaces.

003905 Search

Options

- Starts with any of these keywords
- Starts with the first keyword and contains all of the remaining keywords
- Contains any of these keywords
- Contains all of these keywords

Case insensitive

Choice:

SP-15-003905

Deselect

Cancel < Back Next > Finish

**Figure 18** Prompt page of stand-alone case detail report

## 6.5. Client Tier Implementation

As discussed in Section 5.5, the Client Tier is where the users access the dashboard. Because the case study BI solution uses a web-based user interface, the dashboard is accessible from different devices (supported by the BI tools). Figure 19 and Figure 20 show the overall dashboard view captured from a laptop and from a mobile device. The quality of the visualization and of the interaction is the same in both cases, which helps address the flexibility aspect of the thesis' research problem (Section 1.2).



Figure 19 Pathology dashboard showed on a laptop computer

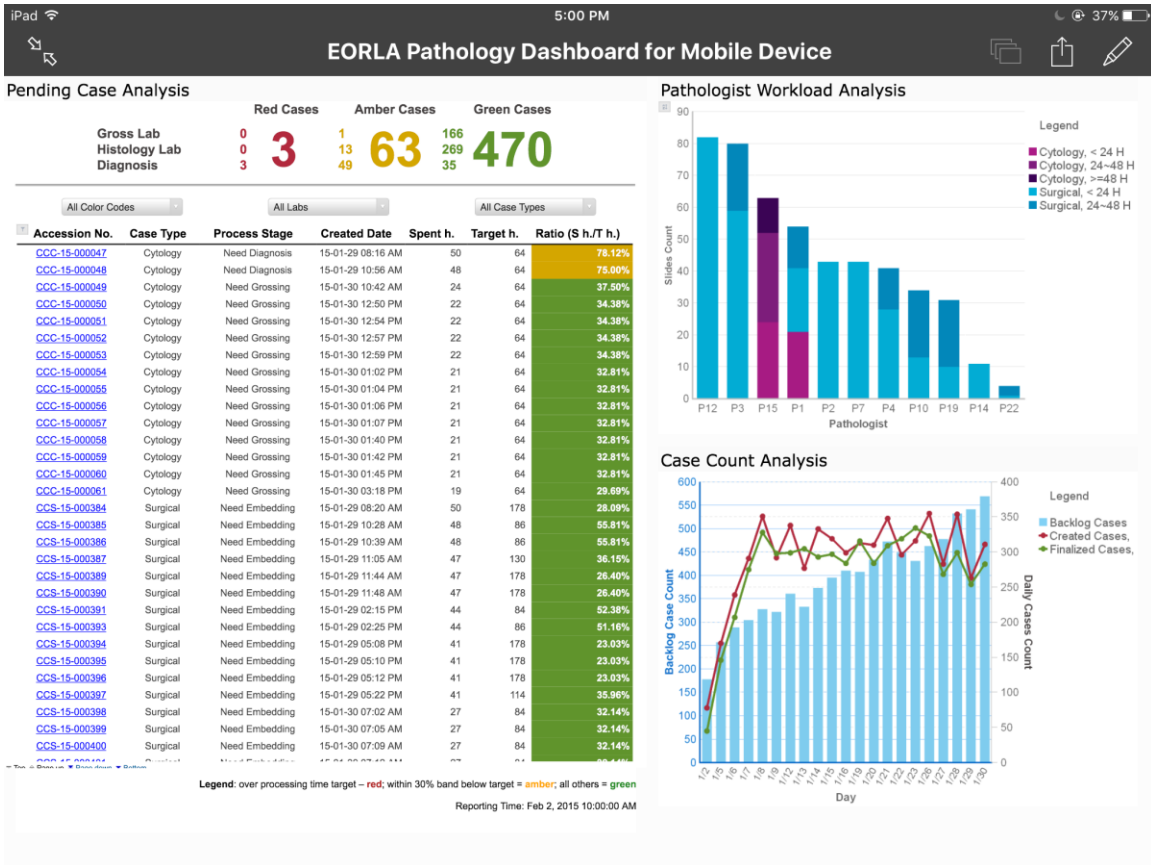


Figure 20 Pathology dashboard showed on an Apple iPad

## 6.6. Enabling Real-time Monitoring in the Dashboard

The literature review in Table 6 shows that most of the pathology dashboard solutions lack the ability to support real-time monitoring. This brings inconvenience to the users because they have to manually rerun or refresh a dashboard in case they want to check the latest data. If a dashboard can refresh by itself, in short time intervals, it can act as a real-time monitoring system, so that end users will always get the latest records without any manual work.

Real-time monitoring was enabled in this case study by adopting the methods discussed below,

- **Disable data cache in report service at the Application Tier:** Using data cache in report service can benefit the performance of some reports, as the reusable query results are cached in memory and thus query request is not sent to the database for each report call [20]. In contrast, data cache needs to be disabled for real-time reporting, in order to force the query request to always run against the database and fetch the freshest data. Local data cache is disabled in this project using the approach introduced in [20].
- **Enable component auto-refresh in dashboard at the Presentation Tier:** Besides disabling data cache at the report service level, one other necessary method to achieve real-time monitoring is to have the dashboard or its components auto-refreshed. To accomplish this, we embedded an HTML item in the dashboard, which includes the JavaScript code shown below<sup>10</sup>. The code in boldface is an adjustable parameter that controls the refresh frequency (in seconds), and system admin can always control the frequency by setting the parameter to a different value. If the `refreshSec` parameter value is set to too low, which means the dashboard will be auto-refreshed at a high frequency, it will bring negative effects to the system, such as (1) underlying database might

---

<sup>10</sup> Additional web browsers configuration may be required to ensure the JavaScript HTML item can work in the Cognos Workspace, for example, Enable JavaScript for the Firefox. For more details, check Cognos Analytics > Cognos Business Intelligence 10.2.2 > Author and Create > Cognos Workspace Advanced User Guide 10.2.2 > Understanding Cognos Workspace Advanced > Web browser settings, [http://www.ibm.com/support/knowledge-center/SSEP7J\\_10.2.2/com.ibm.swg.ba.cognos.ug\\_rptstd\\_fin.10.2.2.doc/c\\_browsers\\_exp.html](http://www.ibm.com/support/knowledge-center/SSEP7J_10.2.2/com.ibm.swg.ba.cognos.ug_rptstd_fin.10.2.2.doc/c_browsers_exp.html)

not able to handle all the query requests, and (2) report service might not been able to refresh and render the dashboard at such high frequency. To avoid the side effects, after discussing with the users, a `minRefreshSec` threshold, initialized to 120 seconds, was added to the JavaScript code. The scripts compare the user's `refreshSec` value with the `minRefreshSec` threshold and use the maximum value as refresh rate.

```
<script language="javascript">
  var func = function() {
    window.oCV_THIS_.executeAction("Refresh", "");
  }
  var refreshSec = 120;
  var minRefreshSec = 120;
  if (refreshSec < minRefreshSec ) {
    refreshSec = minRefreshSec;
  }
  setTimeout(func, refreshSec);
</script>
```

## 6.7. Chapter Summary

In this chapter, we focused on the development work done for each tier of the architecture for the EORLA pathology BI project. Several prototype iterations and design decisions were discussed along the way. At the end of the chapter, we also described how to enable real-time monitoring in the dashboard. Some functions that users can enjoy with the dashboard, such as maximizing the component size and changing the chart display type, are not included in the chapter, because they are provided for free by the underlying BI tool (IBM Cognos here) rather than implemented in the solution. The interested reader can refer to the Cognos manual for a full list of features support by the tool<sup>11</sup>.

The next chapter presents a more formal evaluation of the usability of the dashboard by expert users.

---

<sup>11</sup> Cognos Business Intelligence 10.2.2, Author and Create, Cognos Workspace User Guide 10.2.2, Understanding Cognos Workspace Features available to different capabilities. [http://www.ibm.com/support/knowledgcenter/SSEP7J\\_10.2.2/com.ibm.swg.ba.cognos.ug\\_buxc.10.2.2.doc/c\\_ug\\_exd\\_graduated\\_capabilities.html#ug\\_buxc\\_ibm\\_cognos\\_dashboard](http://www.ibm.com/support/knowledgcenter/SSEP7J_10.2.2/com.ibm.swg.ba.cognos.ug_buxc.10.2.2.doc/c_ug_exd_graduated_capabilities.html#ug_buxc_ibm_cognos_dashboard)

# Chapter 7. Evaluation

---

At the evaluation stage of the DSRM process, in order to evaluate the proposed pathology BI solution, we conducted a usability study at EORLA with six expert participants following a predefined user scenario with different tasks and a questionnaire. In this chapter, we discuss the usability study and analyze the experiment results. The experiment's objective, methods, recruitment, participation, task analysis results, and questionnaire results are discussed in separate sections.

The experiment described in this chapter and the related documents attached in Appendices A to F were reviewed and approved by University of Ottawa's Office of Research and Integrity, Research Ethics Board, certificate H 09-15-05. The experiment's scenario, tasks, and questionnaire were tested informally with two Ph.D. students and one faculty member prior to being used formally with external participants, and feedback and corrections were incorporated in the final version of the experiment.

## 7.1. Objective

The specific objective of this experiment is to validate and refine the solution before pursuing further development. We seek to answer the following research questions:

1. Do people holding managerial positions in pathology perceive the BI dashboard solution as an easy tool to use? This question seeks to evaluate the usability of the solution.
2. Does the BI dashboard solution meet the requirements of people holding managerial positions in pathology? This question seeks to evaluate whether the solution adequately meets users' needs.
3. Is the BI dashboard solution perceived as value-adding by people holding managerial positions in pathology? This question seeks to evaluate if users perceive that the solution is sufficiently helpful to be used in their daily activities.

## 7.2. Methods

This experiment focuses on user feedback in order to improve the customer experience of the dashboard.

To gather user feedback, we followed standard procedures for user interface evaluation [24][46]. Specifically, after a brief introduction to the purpose of the study, each participant was given a short demonstration on how to use this tool. Following the demo, they were also asked to complete a series of tasks with BI tool. This enabled researchers to observe and record data of task duration and about which tasks require additional instructions for different participant; additionally, the researchers took hand-written notes about issues of the tool based on their user experience. Participants were then asked to complete a post-test questionnaire, in which they evaluated the usability of the tool, how well the tool meets their needs, and the potential value they may perceive in using such a tool in their daily activities (see Appendix C – Experiment Protocol, and Appendix F – Questionnaire). Answers regarding the usability of the tool and its ability to meet users’ needs are recorded through a five-point Likert scale, while questions about the perceived value of the tool are recorded through open-ended questions. Participants were also asked about their professional background through multiple-choice questions, and about their experience with any other similar tool through Yes/No and open-ended questions.

Observational data about task duration and the need of participants for additional instructions were compiled quantitatively in order to find averages (e.g., Task 1 required in average 3 minutes to complete; half the participants required additional instructions to complete Task 3). Observational data about emerging issues that participants experienced while using the tool were using basic qualitative content analysis techniques [36]. Specifically, hand-written observational notes were coded in order to find emerging themes; a descriptive matrix was then developed to identify how many participants experienced which issues. This gives the researchers an indication of recurring issues that need to be addressed.

Answers to the questionnaire collected on a Likert scale were aggregated (e.g., 4 out of 6 participants “agreed” or “strongly agreed” that the tool is easy to use). These aggregated results do not relate to a representational sample of the population, and thus cannot serve as the basis for claims regarding the opinions of all pathology managers. However,

these results are useful in pointing out which features or dimensions may be appreciated, and which ones need further development in the next iteration of the tool. Answers to open-ended questions were analyzed in the same manner as the hand-written observational notes.

### **7.3. Participant Recruitment**

Participants included in the experiment are individuals holding management positions in pathology departments or units. We distinguish between operations or laboratory management and clinical management, since the latter type of role refers to clinical activities such as diagnostic rather than daily operations of a pathology facility. Our tool is meant to be used by operations and laboratory managers, not clinical managers. However, since the same person can have more than one type of responsibilities in his/her job, we sought to recruit people holding pathology management roles in general, then we identified their main responsibility (operations, laboratory, or clinical) through the questionnaire that participants were asked to fill out (see Question 18 in Appendix F – Questionnaire). Specifically, we recruited participants within member institutions of EORLA, which is a newly established association of all the laboratory and Pathology departments of Eastern Ontario that currently amalgamates eight hospitals. Reaching out to participants through EORLA gave us access to a pool of experts that could otherwise be difficult to reach.

There are a limited number of potential participants in Eastern Ontario given the small number of pathology facilities in the region. It is unlikely that managers in farther pathology facilities would travel the required distance to participate. While such a limited number does not allow for statistical representation of pathology managers, it is however sufficient for the improvement of the BI dashboard solution given that it stands as a research prototype, not as a tool that would be ready for commercialization.

An email was sent to people holding managerial positions in member institutions of EORLA, informing recipients about the study and inviting them to participate in the experiment (see Appendix A – Contact Email Sent to Potential Participants). The Operations Manager at the Pathology Department of TOH acted as our contact with EORLA and was authorized to send this email on our behalf. This email was sent through the EORLA mailing list by the Operations Manager, but gave the contact information of the Principal

Investigator so that people interested in participating could reply to her rather than to the Operations Manager.

Potential participants answering the email were asked to choose a time slot at their convenience to participate in the experiment. At that point, a copy of the consent letter (see Appendix B – Consent Letter) was also sent to them for perusal; however, they were only asked to sign it at the beginning of their session (see Appendix C – Experiment Protocol). Each session involved only one participant.

The names of participants were only recorded on their consent form. Questionnaires were identified using numbers (i.e., participant 1, 2, 3, etc.). Each participant's number was inscribed on his or her consent form in order to keep a record of the link between participant numbers and names on consent forms, enabling the removal of a participant's data in the case that he or she decided to withdraw from the study at a later date (a situation that did not happen in the end). The consent forms identifying names and participant numbers are kept in a locked cupboard located in a locked laboratory room at the Telfer School of Management. Questionnaires are kept in a separate locked cupboard in the same locked laboratory room. Only Professor Michalowski (one of the co-supervisors) has a key for each cupboard.

We were aiming to recruit between four and eight participants for the experiment, and were able to recruit six participants with various roles from EORLA. Two of the participants had seen the solution demo previously, and the other four had no knowledge whatsoever about this project.

## **7.4. Participation**

### **7.4.1 Overall Description**

Experiments were conducted with only one participant at a time. A list of tasks was provided to the participant, who was asked to conduct these tasks in order to answer specific questions using the BI dashboard tool (see Appendix D – Participant Tasks). Since the goal of a session is to gather feedback rather than evaluate participants' abilities in using the dashboard, each participant was provided with help to accomplish these tasks as needed (see Appendix E – Task Answers and Instructions). This activity lasted about 30 minutes.

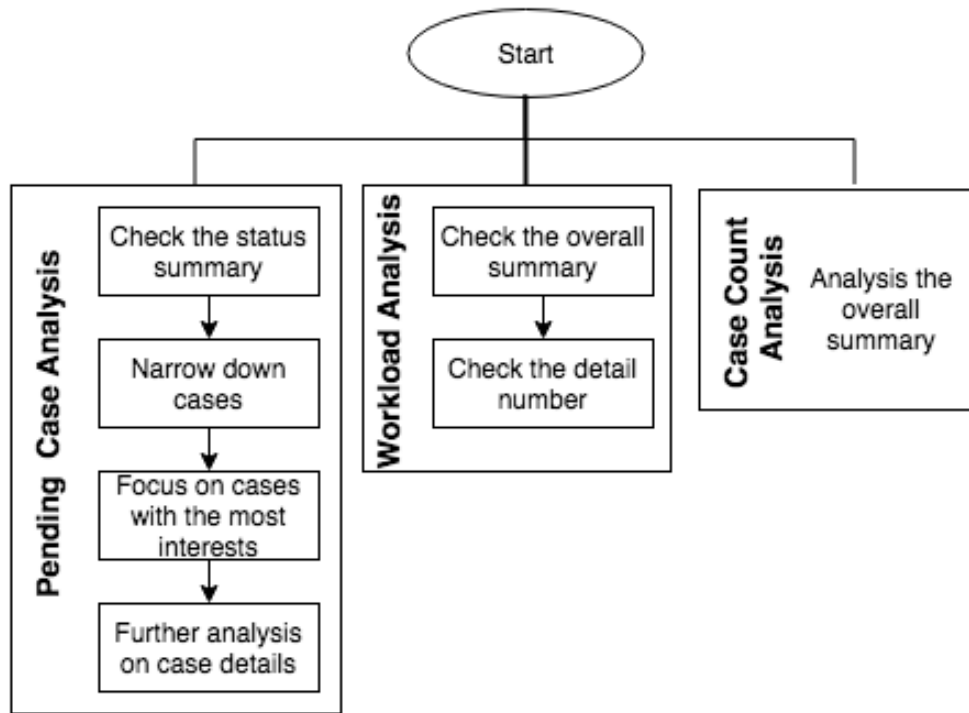
After that, the participant was provided with a questionnaire intended to answer the research questions (see Appendix F – Questionnaire). Questions 1 to 4 have been taken from an earlier, similar study aiming to evaluate the usability of an online tool [32]. Questions 5 to 14 are directly derived from the requirements used to develop the tool. Questions 16 to 18 aim to assess whether participants perceive the proposed tool as valuable in the context of their work, and how it could be improved to do so. Questions 19 and 20 are contextual questions. Question 19 seeks to identify the type of managerial activities the participant is involved in, and Question 20 seeks to know if the participant is familiar with BI dashboard tools. Answering the questionnaire took about 15 minutes.

#### **7.4.2 Predefined User Scenario**

The list of tasks in the experiment follows a predefined user scenario: *as a management staff from the pathology department, how could the BI dashboard help me if I want to figure out the answers to my everyday department/lab management questions?*

Figure 21 illustrates this predefined user scenario. The tasks start with Pending Case Analysis. Tasks 1 – 4 guide the participants to read and interact with this component. Task 1 requires participants to focus on the upper static section of the component and understand the meaning of font size and colour code. When it comes to Task 2, participants were asked to use the dynamic value prompts group to narrow down cases in the detail list. Task 3 requires participants to sort the detail list, and Task 4 leads participants to trigger the drill-through behaviour of the case detail analysis. Task 5 focuses on the case detail report, in which the participants are required to read the case details to complete the task.

Tasks 6 and 7 are designed for the Pathologist Workload Analysis. Task 6 asks the participants to observe the stacked bar chart to obtain the overall summary, while Task 7 requires them to convert the chart to a table in order to get the actual numbers. Task 8 is an open-ended question on the Case Count Analysis, which requires participants to analyze the relationships in the created cases, finalized cases, and backlog cases. This list of tasks aims to guide the participants to walk through the pathologist BI dashboard daily process and experience the major features of the dashboard. The solution requirements discussed from Section 3.2 are also covered in this task list.



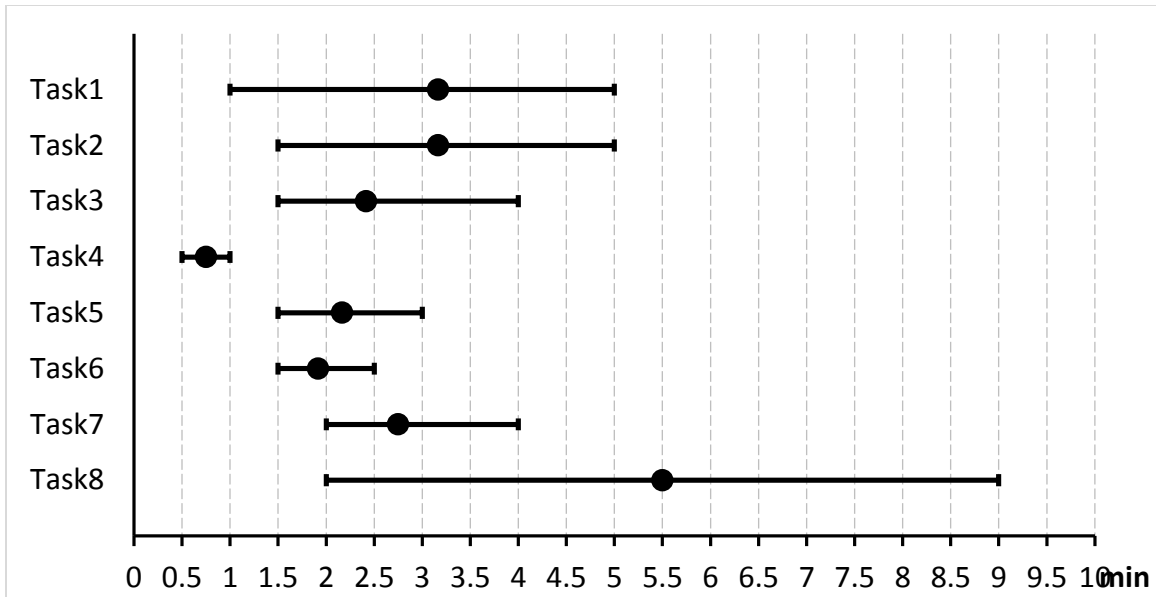
**Figure 21** Predefined user scenario

## 7.5. Task Result Analysis

### 7.5.1 Timing Analysis

The first measurement we took from the experiment is the time that a participant spent on an individual task. This measurement tells us how efficient the tool is to help a user find out the answer to a question. Because of the facts that (1) only two participants had seen the solution previously whereas the other four had never seen the demo, and (2) it was the first time participants had to use the tool just after a 10-minute demo, the measurement are good indicators of whether the solution is easy to learn and easy to use.

The timing interval was set to 0.5 minute during the data recording. The minimum, average, and maximum times taken for each task are shown in Figure 22.



**Figure 22** Task time measurement for all participants (minimum, average, maximum)

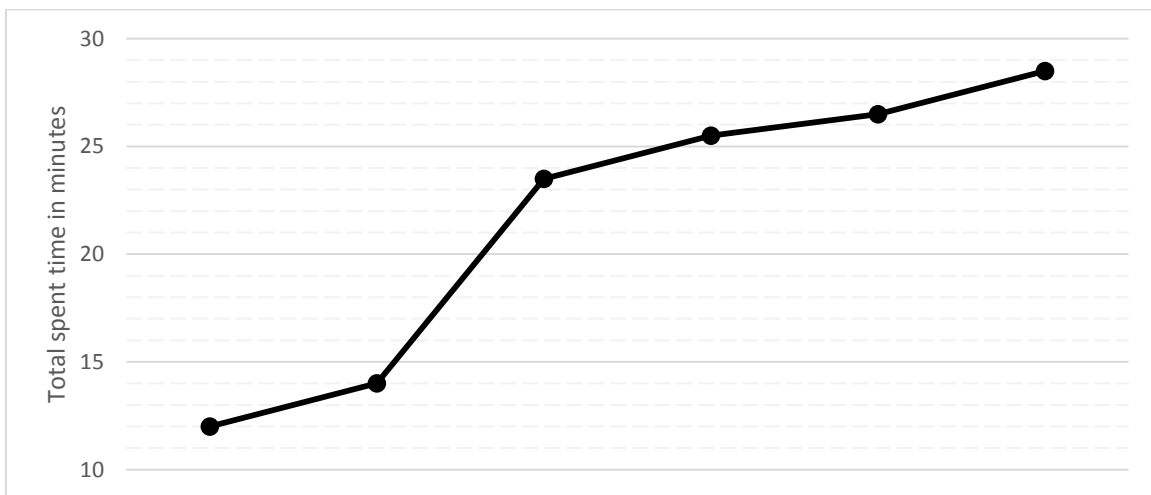
From the above figure, we find that among the questions related to Pending Case Analysis (Task1 to Task5), Task1 “Check the status summary” and Task2 “Narrow down cases” take a relatively long time compared to Task3 to Task5. The reasons could be:

1. These two tasks are at the beginning of the experiment and participants took some time to get familiar with the tool, such as what the font size and colour code mean in the component;
2. Tool performance with user interactions. For example, the value prompts in the prompt group are set to auto-submit. When a user clicks an option from the dropdown list, the value prompt directly submits the selected value to the server end and then refreshes the case detail list with filtering on the selected value. Task2 requires a participant to interact with all the value prompts, which will trigger this “auto-submit & refresh” behaviour three times, and it becomes costly for the tool to complete such tasks in terms of processing time;
3. Potential design flaw. Finishing Task1 is supposed to take little time as all the numbers can be read directly from the static section. One possible explanation as to why this is taking longer than expected for some participants is that the design itself may cause difficulties for a user to understand the component.

The timing results related to Pathologist Workload Analysis (Task6 and Task7) are as expected, since Task6 requires a reasonable time to analyze the colours and the stacker bar, and Task7 requires user interactions on the chart type conversion.

Task8 is an open-ended question on the Case Count Analysis, and we observed a large variation of the times spent on this task. Some participants from the management of the department who had seen an earlier prototype answered Task8 quickly, whereas other participants from the lab management required more time to analyze the correlation between the three measurements included in the component (created cases, finalized cases, and backlog cases). Hence, either using the tool more often helps users perform such tasks quickly, or improvements are needed for one of the target groups of users (lab managers).

Figure 23 shows the total spent time by each participant, sorted from the lowest to the highest value. The data points fall into two distinct ranges: one is within a 10–15 minute range, and the other is within a 23–28 minutes range. Participants who had previously seen the solution demo are in the first range and spent much less time than the “inexperienced” participants. Thus, more demo and guidance about the solution will likely help flatten the learning curve, and we expect experienced users to be able to perform these representative tasks in a short time. Checking this, however, is outside the scope of this experiment.



**Figure 23** Total spent time in minutes, sorted in increasing order

## 7.5.2 Correctness Analysis

### Grading Criteria

One important purpose of this experiment is to verify whether the BI dashboard can help the participants get correct answers. The following answer grading scale was developed to judge task answers:

- **Correct:** A task and the written answers for the task questions are correctly completed by a participant;
- **Partial:** A task or the written answers for the task questions are partially completed by a participant;
- **Error:** A task and the written answers for the task questions are incorrectly completed by a participant.

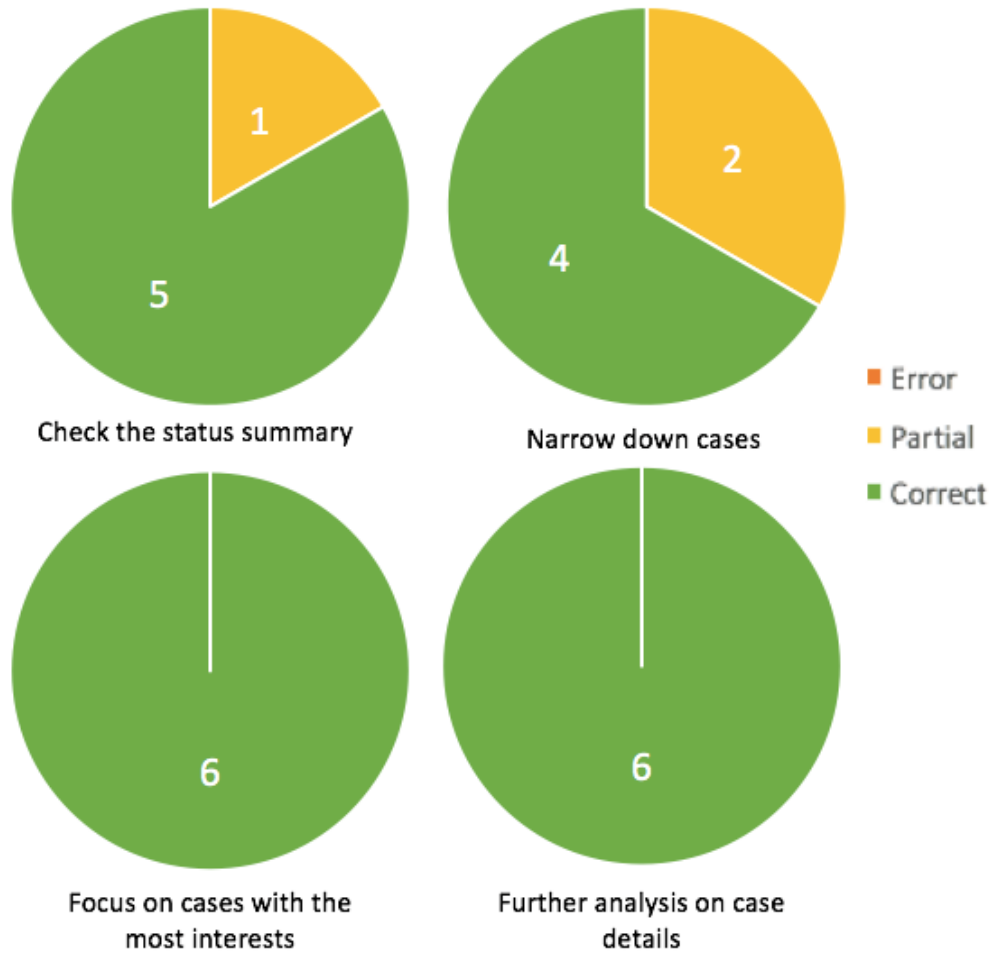
### Correctness of the Pending Case Analysis

Figure 24 illustrates the correctness of tasks related to the Pending Case Analysis. As shown in the pie chart for Task1 - “Check the status summary”, five out of the six participants have completed the designated tasks with fully correct answers while only one participant was partially correct. For Task2 - “Narrow down cases”, four were correct and two were partial. For both “Focus on cases with the most interests” and “Further analysis on case details”, all of the participants had them successfully finished.

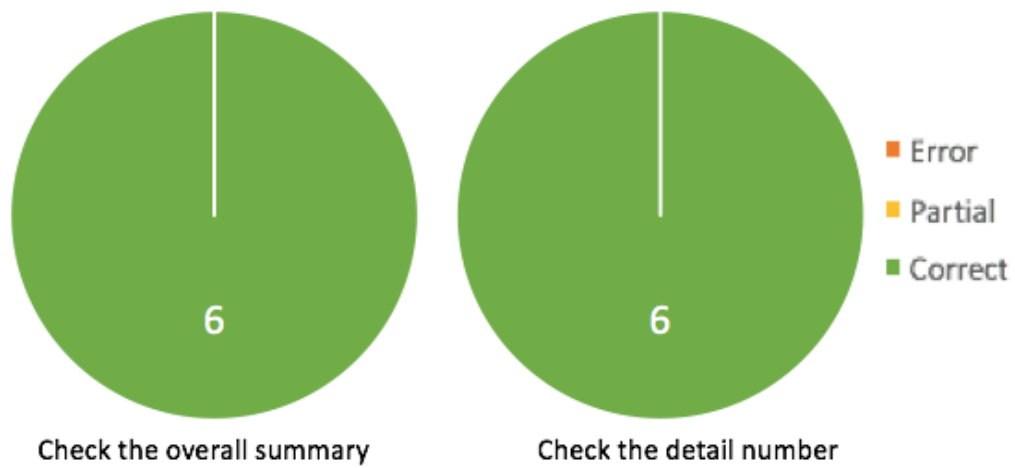
While participants were doing tasks, we noticed that some participants were hesitant on Task1 and Task2, and we also found user behaviours that would lead to the incorrect answers. For example, a participant neglected the static section of the component and tried to get the answer of Task1 from the case detail list. On Task2, some participants selected incorrect options from the dropdown list value prompts. Oral guidance was provided for situations like these. For Task3 to Task5, participants ended up working with more confidence, and no assistance was provided.

### Correctness of the Pathologist Workload Analysis

The correctness of tasks related to the Pathologist Workload Analysis is shown in Figure 25. We received great correctness results for this component. Participants finished related tasks with strong confidence and no participant asked for additional help.



**Figure 24** Pending Case Analysis correctness pie charts



**Figure 25** Pathologist Workload Analysis correctness pie charts

### Correctness of the Case Count Analysis

The correctness of the Case Count Analysis is illustrated in Figure 26. A significant discrepancy of performance on this open-end question was found between participants from different management roles. It seems relatively easier for participants from department management to derive a thorough answer whereas participants from lab management struggled when answering the same question.

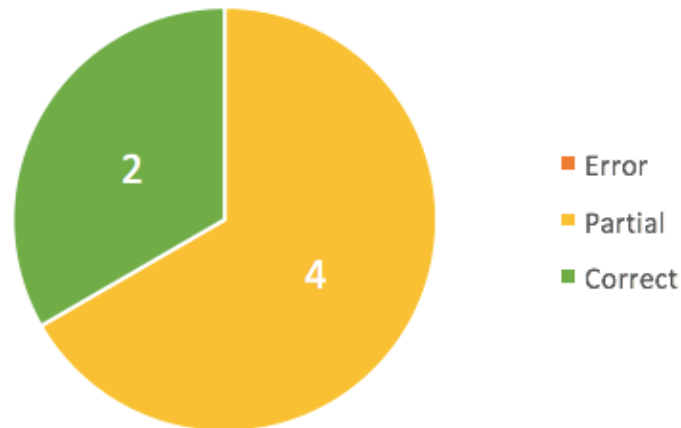


Figure 26 Case Count Analysis correctness pie chart

### 7.5.3 Multi-dimensional Analysis

#### Multi-dimensional Analysis for Participants

Regarding the multi-dimensional analysis of results for participants, we developed a weighted score scale to show how successful a participant is when completing a list of tasks. The denominator of the formula comes from the total number of tasks, with a maximum of 2 points each. This score is applied to each participant of the usability study:

*Weighted Score*

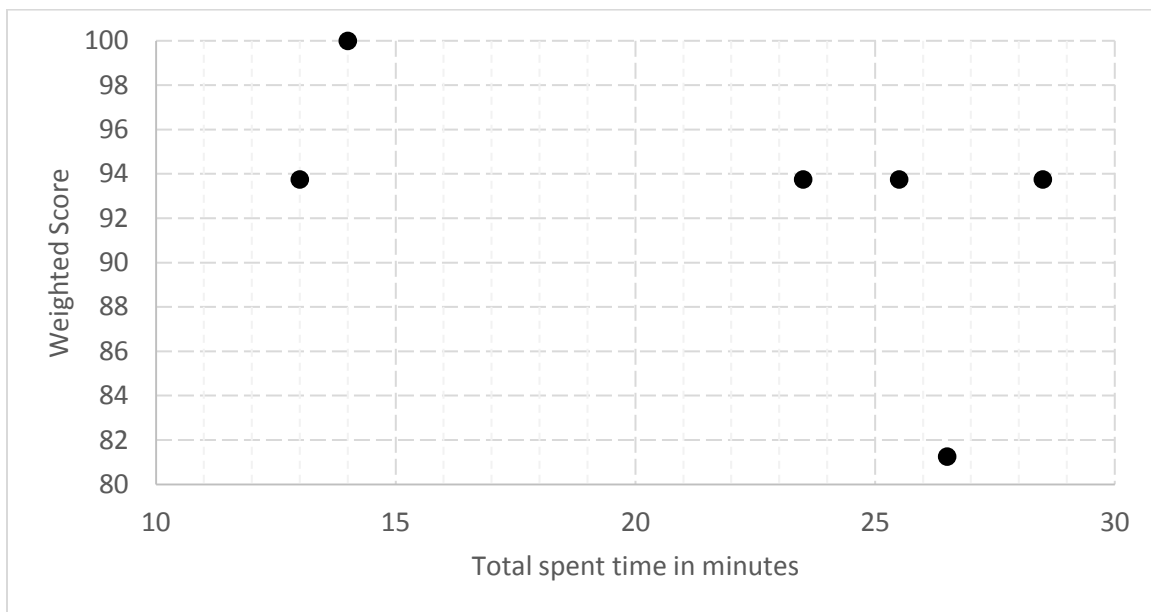
$$= \frac{(2 \times \text{Correct} + 1 \times \text{Partial} + 0 \times \text{Error})}{2 \times \text{TotalNumberOfTasks}} \times 100$$

After calculating the weighted scores, we made a scatter chart using the total spent time and the weighted score of each participant, as shown in Figure 27. The distribution of the number points indicates that:

1. All participants finished the list of tasks with a high degree of completion. Based on the weighted score formula, the range of the score is from 0 to 100.

As observed, five out of the six participants received a score of 93.75 or higher, and only one of the participants got a score in the 80–90 range. This result suggests that the participants, as first-time users, can finish the list of tasks using the proposed dashboard without a steep learning curve.

2. The score of each participant does not closely correlate to the time spent on the tasks. Even though the total spent time by participants divides into two groups (below 15 minutes and above 23 minutes), this difference is hardly impacting the correctness score.



**Figure 27** Cross analysis by participant scatter chart

### Multi-dimensional Analysis by Tasks

In order to conduct a multi-dimensional analysis by tasks, Table 12 was created to present the total spent time, the correctness level, and the guidance count for each task. This table indicates one potential relationship between these three variables: when a participant takes a relatively long time to finish a task, the correctness of the task completion decreases, and the participant tends to require more assistance on the task. Potential explanations could include (1) the difficulty of the tasks varies, and (2) the design and implementation of the solution have some flaws and need further improvement. A discussion about potential improvements is included in the next chapter.

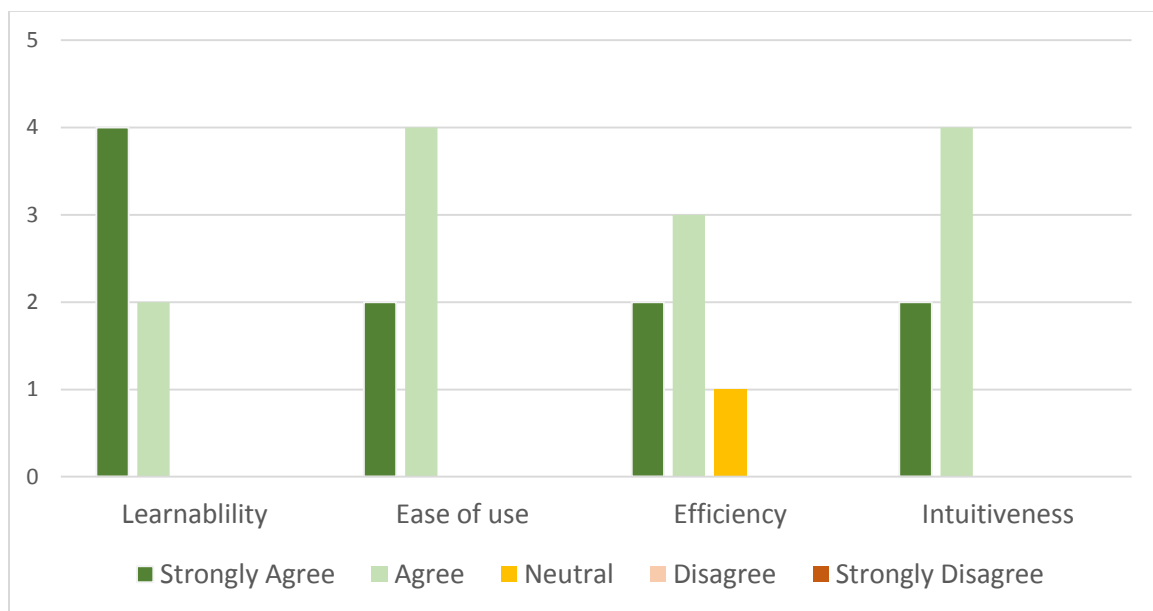
**Table 12** Multi-dimensional analysis by task result table

	Total Spent Time (in minute)			Correctness			Guidance Count
	Min.	Max.	Avg.	Correct	Partial	Error	
<b>Task1</b>	1	5	3.17	5	1	0	2
<b>Task2</b>	1.5	5	3.17	4	2	0	2
<b>Task3</b>	1.5	4	2.42	6	0	0	0
<b>Task4</b>	0.5	1	0.75	6	0	0	0
<b>Task5</b>	1.5	3	2.17	6	0	0	0
<b>Task6</b>	1.5	2.5	1.92	6	0	0	0
<b>Task7</b>	2	4	2.75	6	0	0	0
<b>Task8</b>	2	9	5.50	2	4	0	6

## 7.6. Questionnaire

At the end of the experiment, participants were asked to complete a questionnaire intended to help us understand how the participants feel about the solution.

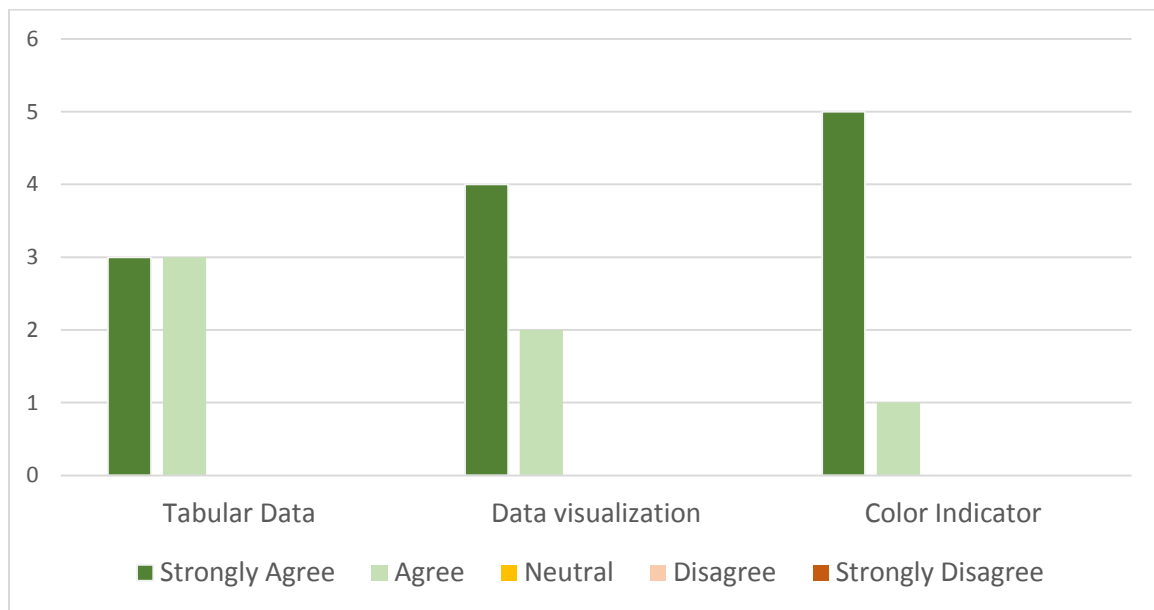
The first part of the questionnaire is about the overall impression of the “use and learn” experience, and four related questions were included.



**Figure 28** Questionnaire results on the “use and learn” experience

The results in Figure 28 indicate the participants agreed that the proposed tool is easy to learn and use, and that it is easy to understand the output presented on the user interface (intuitiveness). Most of the participants considered that the solution can help them perform tasks more efficiently.

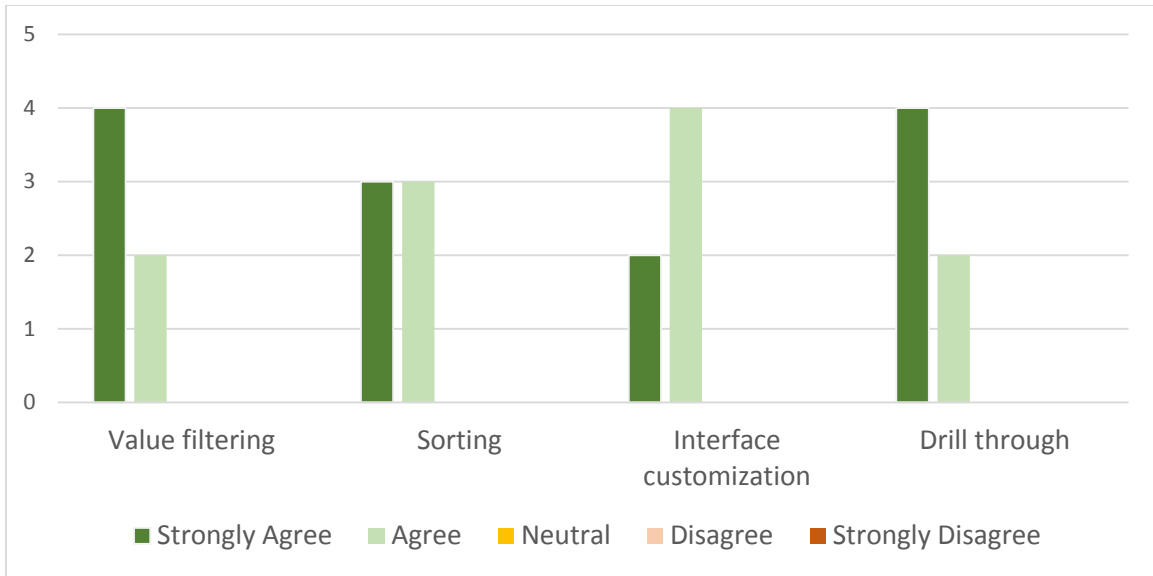
Figure 29 illustrates the questionnaire results in terms of UI presentation of the pathology BI dashboard. While all participants felt the proposed tool is efficient in terms of data presentation using tabular data and data visualization, they were also satisfied and strongly in favor of the colour indicators being used in the dashboard.



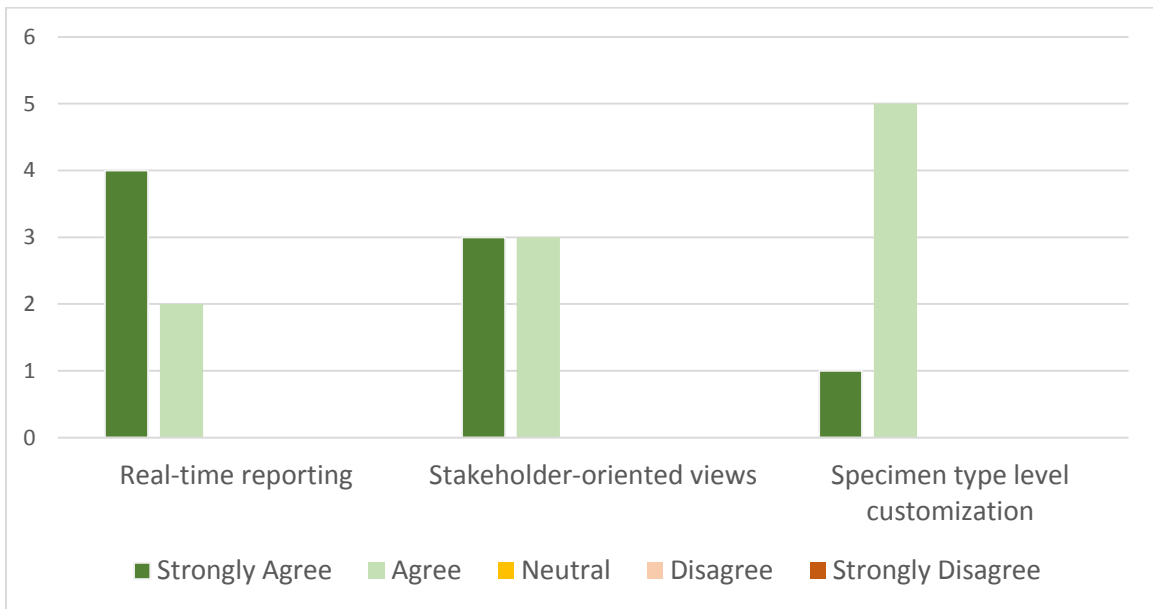
**Figure 29** Questionnaire results on UI presentation

In addition, Figure 30 shows great satisfaction about the user interaction experience of the dashboard. Interactions like value filtering and drilling-through have a high user satisfaction, while sorting and interface customization have a good satisfaction.

Questionnaire results concerning the advanced features are shown in Figure 31. We found that participants considered the proposed tool to be very good at real-time reporting and at managing stakeholder-oriented views. Although the satisfaction level is good for specimen type level customization, there might be an opportunity here for improvement.



**Figure 30** Questionnaire results on user interaction

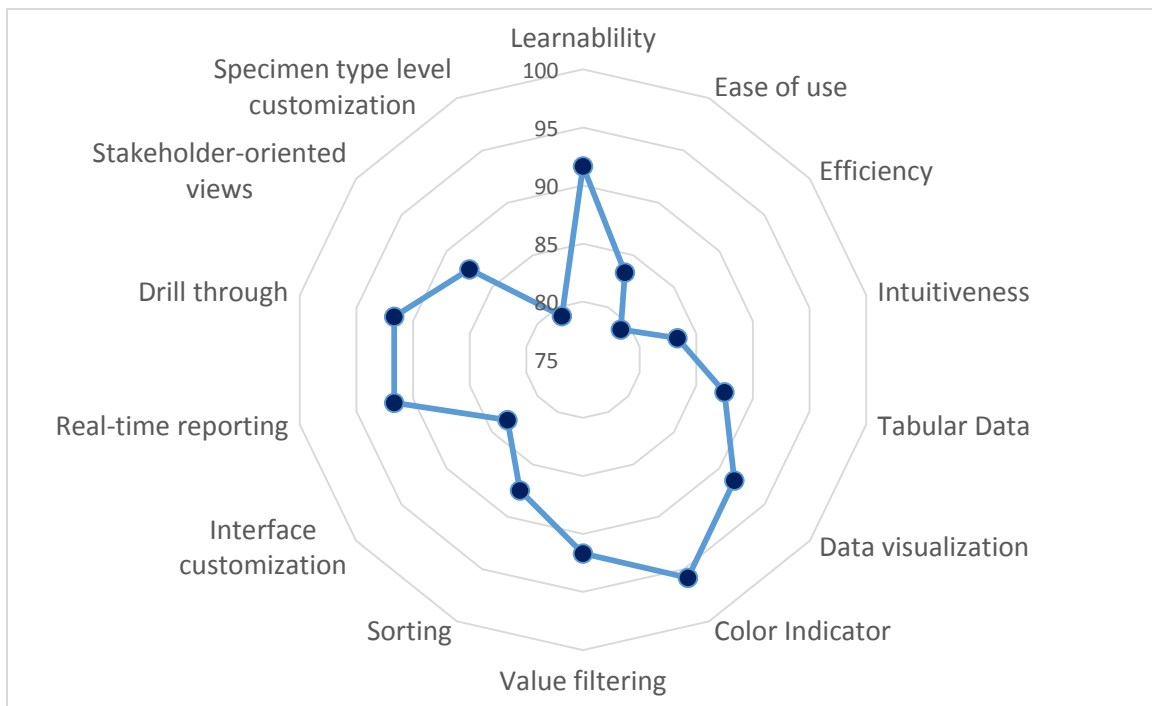


**Figure 31** Questionnaire results on advanced features

In order to analyze the aggregated user satisfaction of the proposed solution, we developed a weighted score scale to calculate the satisfaction score for Questions 1 to 14 in the questionnaire. The denominator of the formula comes from the total number of participants, with a maximum of 4 points each. This score is applied to each of the 14 first questions of the questionnaire:

$$\text{Satisfaction Score} = \frac{4 \times \text{Stongly Agree} + 3 \times \text{Agree} + 2 \times \text{Neutral} + 1 \times \text{Disagree} + 0 \times \text{Strongly Disagree}}{4 \times \text{TotalNumberOfPaticipants}} \times 100$$

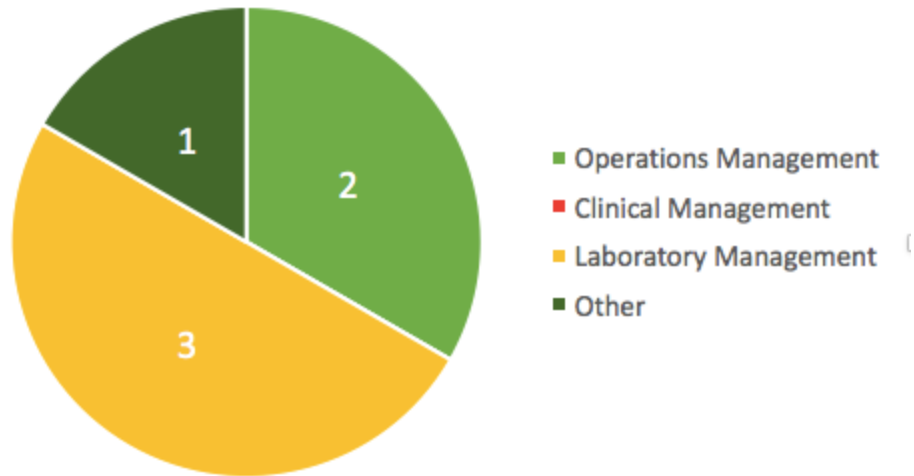
The radar chart in Figure 32 shows the aggregated satisfaction scores for the 14 different aspects of the proposed BI solution surveyed in the questionnaire. We observe here that “Colour Indictor” achieved the highest score (95+). Most of the satisfaction scores fall into a good 80–95 range, and two items (“Efficiency” and “Specimen type level customization”) had slightly lower scores (75–80) compared to the others. We considered the solution received a high overall satisfaction from all the participants, as all the aggregated satisfaction scores are higher than 79.



**Figure 32** Total satisfaction score radar chart

The questionnaire also included three open-end questions, which aimed to know whether participants would consider the proposed BI solution as a valuable tool in their work environments, with a rationale, and what suggestions the participants have to the proposed tool. A discussion of these open-end questions is included in the next two chapters.

At the end of the questionnaire, we sought to identify the managerial roles of the participants. The distribution is shown in Figure 33. We also asked the participants whether they were familiar with any BI dashboard tools, and all of them answered “No”.



**Figure 33** Questionnaire results on management role

## 7.7. Chapter Summary

In this chapter, we described how the evaluation of the proposed BI solution was performed. Through the analysis of the results of pre-defined tasks in terms of time and correctness, and the analysis of the perceived quality and usefulness via a questionnaire, we concluded that the proposed dashboard is of good quality and led to a high level of satisfaction among the six expert participants.

The next chapter compares the proposed dashboard with closely-related work, reviews the suggestions from the questionnaire's open-ended questions, and discusses limitations and potential threats to the validity of this study.

## Chapter 8. Discussion

---

This chapter compares the dashboard developed as a result of thesis research with those discussed in Chapter 4 considering the stakeholder requirements from Chapter 3. User feedback on the dashboard is also discussed to provide another viewpoint on the research success. Limitations and threats to the validity of this study are finally evaluated at the end of this chapter.

### 8.1. Technical Comparison

To assess the technical completeness of the dashboard developed in this thesis, we check it against the requirements from Chapter 3. These twelve requirements, supplemented by two categorizations, were already used as criteria for the related work in Table 6.

The proposed dashboard is of *tactical and operational* type as its typical users are pathology department managers and lab operational managers. The chart types used in the new dashboard are the same as for existing solutions, because they are the common way to present trends by time (line chart) and show comparisons (column chart).

The new dashboard meets the stakeholder requirements related to user interfaces. It fulfills the *Tabular Data* and *Data Visualization* requirements by listing detailed data in tables (in the Pending Case Analysis component and in the Case Detail report) and by presenting summarized data in charts (in the Pathologist Workload Analysis component and in the Case Count Analysis). The *Colour Indicators* requirement was implemented in all three components in the dashboard, using metaphors such as a “traffic light” colour code to show different case status in the Pending Case Analysis component, and a “heat map” colour palette to differentiate case types and case spent time in the Pathologist Workload Analysis component.

The new dashboard also meets the stakeholder requirements related to user interactions. The *Value Filtering* requirement is fulfilled by the value prompt group in the Pending Case Analysis, which allows users to dynamically filter the case list by a combination of selected colour code, lab, and case type. The *Sorting* requirement is satisfied in the case

list as well, and users can sort by ascending or descending order using any column in the list. *Drilling-through* is enabled on case accession numbers, which helps users closely examine the details of a case. The implementation of the *Interface Customization* requirement is available in the Pathologist Workload Analysis, where it is possible for the users to convert the chart to a table and check the actual numbers behind the graph.

Three out of five stakeholder requirements on advanced features are implemented in our new dashboard. To satisfy the *Real-time Reporting* requirement, JavaScript code was embedded in the Pending Case Analysis component to trigger the dashboard in refreshing itself in user-defined time interval, for example 120 seconds. *Stakeholder-oriented View* is fulfilled by allowing managers from different organization levels, such as department managers or lab managers, to quickly focus on their respective business view in the Pending Case Analysis. The *Specimen-type Customization* requirement is met in the backend calculation, which applies different business rules per specimen type to calculate the case target time.

*Automatic Notifications* should be simple to implement (as Cognos BI product supports watch rules that can trigger email notifications) but they are not yet supported in the dashboard as more analysis is needed to define appropriate rules and notification destinations. Finally, *Workload Planning* requires much more investigation as to whether useful predictions can be obtained by modelling the available data, and hence this requirement remains outside the scope of the current dashboard version.

Table 13 summarizes the evaluation of the new dashboard implementation against the requirements and compares it to selected similar solutions. These solutions (initially listed in Table 6) concern the generic healthcare dashboard with the largest number of satisfied requirements (Wadsworth et al. [50]), the best two pathology-specific dashboards (Viewics Inc. [4] and Kofax [3]), and the PowerJ prototype [16] developed at the DPLM. The dashboard developed in this thesis (last line of Table 13) meets the largest number of requirements overall and particularly excels in real-time reporting and stakeholder-oriented views. There is room for improvement in terms of support for automatic notifications and workload planning, but none of the other dashboards meets these requirements either.

**Table 13** Comparison of the thesis with the closely-related work

Article/Product	Dashboard Type	Chart Type	Tabular Data	Data Visualization	Value Filtering	Real-time Reporting	Colour Indicators	Sorting	Drill Through	Interface Customization	Stakeholder-oriented Views	Specimen-type Customization	Automatic Notifications	Workload Planning
Wadsworth et al. [50]	Strategic Tactical Operational	Line chart Column chart	Y	Y	Y	N	Y	N	Y	N	N	N	N	N
Viewics Inc. [4]	Tactical	Line chart Column chart	Y	Y	N	+/ -	Y	N	N	N	N	N	N	N
Kofax [3]	Tactical Operational	Line chart Column chart Bar chart	Y	Y	Y	+/ -	Y	Y	Y	Y	N	Y	N	N
PowerJ [16]	Tactical	Line chart Column chart	Y	Y	N	+/ -	+/ -	N	N	N	N	N	N	N
This thesis	Tactical Operational	Line chart Column chart	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N

## 8.2. User Feedback on Existing Features

In addition to the analysis of the new dashboard design and implementation from a technical perspective, we also need to understand how users feel about the dashboard. The usability study discussed in Chapter 7 provides necessary information.

### 8.2.1 User Satisfaction about the Stakeholder Requirement Fulfillment

Questions 5 to 14 in the usability study questionnaire (Appendix F) were designed to understand user satisfaction on how the proposed dashboard fulfills the stakeholder requirements. All of these questions have the same format:

*“[Definition of a stakeholder requirement]. Given this definition, would you say the proposed tool provides useful support for [the stakeholder requirement]?”*

Related results are presented in Figure 29, Figure 30, Figure 31, and Figure 32. All of the 60 answers were “Strongly Agree” or “Agree”, with the exception of one participant responding “Neutral”. As a conclusion, we claim that the expert users who participated the usability study agree that the proposed dashboard meets the stakeholder requirements, except for *Automatic Notifications* and *Workload Planning*, which were not included in the experiment. In addition, all the participants would like to use the proposed tool in its current state in the context of their daily professional activities (Question 16 in the questionnaire).

### 8.2.2 User Feedback on Decision Making

Two open-ended questions (#15, #17) asked users about which types of decisions, if any, the proposed tool could help them make, and how users would use the proposed tool. Answers to these questions can be grouped as follows:

- **Bottleneck detection.** “Where are the bottlenecks in the processes which are causing backlog?” and “Make sure we are addressing the cases that are red flagged.”
- **Early alarm.** “Early corrective actions could be taken to avoid backlogs.” and “Ensure that all patient diagnoses are available when required by clinicians.”
- **Turn around time improvement.** “Aid in determining oldest cases and help to decrease turn around time in lab.”, “Determine average turn around time.”, and “Find cases that are over the turn around time and focus on getting these cases updated.”

In summary, the respondents felt the new dashboard is capable of assisting them in making decisions to smooth the pathology facility work process, which should result in fewer bottlenecks, smaller backlogs, and shorter turnaround times.

### 8.3. Limitations and Threats to Validity

Amongst the various types of threats to validity that exist [10], there are three relevant types that need to be discussed in the context of this thesis. They usually reflect limitations of the dashboard or of its evaluation.

- **Construct validity.** As discussed in Section 1.2, the research objective of this thesis is whether “*a real-time business intelligence solution exploiting barcode tracking data can provide efficient, flexible, and multi-tenancy support to pathology department management*”. The major threat to construct validity is that, without an actual deployment of the solution in pathology facility and a certain period of time of its daily usage for the evaluation, it is very difficult to provide data to show its effectiveness and usefulness in supporting the management. The user-based evaluation reported here is limited to usability and *perceived* usefulness.
- **Internal validity.** A few factors in this thesis research may introduce bias threatening internal validity. The first one is about the selection of related studies and solutions. Though guided by her supervisors, the thesis author is the only person who performed the literature review, and hence may have been biased in doing so. Another threat pertains to the use of the stakeholder requirements to evaluate related work. As different projects have different user requirements, using just the EORLA stakeholder requirements as the evaluation criteria might introduce bias. The third factor is the BI tool selection during the solution implementation, which can cause bias on learnability and ease of use when evaluating the proposed dashboard; it is difficult to separate the dashboard design from its underlying implementation technology. The fourth threat to internal validity is the limited number of participants in the usability study, especially as they were recruited from single organization. The fifth threat pertains to the “real-time” technique that we implemented in this solution, which can result in slight differences between data shown on the dashboard and data storing in the database. Such differences can be caused by data refresh delay (up to 2 minutes as per the default dashboard refresh frequency). The sixth threat is that the database structure of the EORLA production database was unknown to the solution design and implementation, thus the restored database and the data model of the solution might have conflicts with the production database on table constraints and join relationships. This last issue may require further investigation at deployment time.

- **External validity.** The first threat to the external validity relates to whether the proposed solution is generalizable to other processes and organizations, but also to other BI tools. The Cognos BI tools used in the thesis study might not be the ones that EOLRA will chose for its production usage, and this may have an impact on the usability and usefulness. Whether the thesis results can be easily adopted using other tools is a threat. In addition, the dashboard was developed to meet EORLA's requirements, and may not work well in other organizations with slightly different roles, processes, technologies, or challenges. These are potential threats but can be easily addressed. Other threats to general adoptability might be caused by performance, security, and privacy issues not considered in the thesis.

#### **8.4. Chapter Summary**

In this chapter, we first compared the thesis' dashboard with similar ones by using the stakeholder requirements as comparison metric, and concluded that the new dashboard best met these requirements. Then, we discussed the results of user satisfaction study through the analysis of many questions from the usability study questionnaire, and claimed that the proposed tool received high satisfaction and is perceived as efficient in helping users make decision. At the end of this chapter, important threats to construct validity, internal validity, and external validity were identified. Some of these threats and associated dashboard/evaluation limitation help to identify relevant future work, which is the subject of the next section, together with a general conclusion.

## Chapter 9. Conclusions and Future Work

---

In this chapter we provide a general conclusion related to the research problem and the thesis objective, and then we list future work opportunities found during the research and the usability study.

### 9.1. Contributions

In order to meet the thesis objective discussed in Section 1.2 – *a real-time business intelligence solution exploiting barcode tracking data can provide efficient, flexible, and multi-tenancy support to pathology department management* – we started by reviewing the closely related academic research and industry solutions to understand the domain and find gaps in existing work. A case study for the EORLA pathology BI project was conducted. We interviewed stakeholders to understand the business context of the BI project, including the business drivers, business goals and stakeholder requirements, and business strategies. A high-level architecture of the case study was proposed based on the general BI abstract architecture, and this solution-specific architecture guided the design and implementation of the proposed solution. After a fully functional solution was developed, a usability study was performed to evaluate it. We also compared the proposed tool with closely related work against the stakeholder requirements.

From both quantitative measures (Section 8.1) and qualitative questionnaires (Chapter 7 and Section 8.2), the proposed BI solution is shown to outperform the other solutions (Table 6 and Table 13) while achieving the research objective.

The major research contributions of this thesis research include,

- Further confirmed and demonstrated that BI solution is an efficient way to help improve an organization management, discussed in Chapter 7 and Section 8.2.
- Established a set of generalized stakeholder requirements that can be used in any pathology management BI solution design and implementation, discussed in Chapter 3.

- Proposed a BI dashboard layout design with a combination of colour indicators, tables and diagrams for the pathology workload management, discussed in Section 5.4 and as shown in Figure 19, Figure 20;
- Proposed a low-cost method to enable real-time monitoring in BI dashboard, discussed in Section 6.6.

The major practical contributions of this thesis include the implementation of a four-tier BI architect solution (Chapter 5 and Chapter 6) and the usability study to evaluate the dashboard (Chapter 7). The proposed dashboard has these major features:

- UI presentation, using suitable data presentations (by tables or graphs) and colour indicators to translate and summarize the barcode tracking transaction records to formats that are efficient in information communication and user understanding;
- User interactions, supporting different types of mechanisms in the dashboard, such as value filtering and sorting, to allow users to customize the information displayed in the dashboard and quickly focus on their data/KPIs of interest;
- Advanced features, enabling the auto-refresh in the dashboard component to make the dashboard provide a real-time monitoring and alarm system, and providing stakeholder-oriented views to enable multi-tenancy support for dashboard users in different user groups.

## 9.2. Future Work

A number of new research opportunities are identified based on the research presented in previous chapters and considering the threats identified in Section 8.3.

1. The Automatic Notifications and Workload Planning requirements are not well supported by the proposed dashboard. The design and development of these two features would improve the dashboard and benefit users managing pathology work processes.
2. There are many potential usability improvements that derive from the usability study (Question 18 in the questionnaire):

- There is an opportunity to adapt the Stakeholder-oriented View in the Case Count Analysis component, so that laboratory managers can gain better tailored insights.
  - The dashboard could support a drill-through functionality in the Pathologist Workload Analysis and Case Count Analysis components, which will allow users to have a closer look at details of the pending cases of a pathologist (Pathologist Workload Analysis) and at the created, finalized, and backlog cases of a day (Case Count Analysis).
  - The dashboard could provide value prompts in Case Count Analysis, so that users can self-define the start date and end date of the analyzing window.
  - The dashboard could adopt the OLAP approach (see Table 4) to enable a multi-level date view in Case Count Analysis, which would allow users to analyze case counts over time using a date hierarchy, such as daily, weekly, monthly, and quarterly.
  - The dashboard could include a mechanism to easily edit business rules, as current approach involves hard coding of these rules.
3. The dashboard should be deployed and tested *in vivo* over a long period of time in order to assess its impact.
  4. The dashboard could be adapted to other similar facilities and re-evaluated (ideally with more participants, although the number of managers is usually low) in order to better assess its suitability in a larger context, beyond EORLA's.
  5. The dashboard could use a sliding window query process over streaming data [15] to resolve the real-time querying limitation discussed in Section 8.3.
  6. Other non-functional properties, including performance, security, and privacy should also be evaluated.

## References

---

- [1] Abela, A. “Chapter 7: Visual Presentation Elements - Graphics, Charts, Colour, Animation, and Fonts”. *Advanced presentations by design: Creating communication that drives action*. John Wiley & Sons (2008)
- [2] Adler-Milstein, J., DesRoches, C.M., Jha, A.K.: “Health information exchange among US hospitals”. *The American journal of managed care*, Vol. 17, Issue 11, pp. 761–768 (2011)
- [3] *Anatomic Pathology Dashboard*, Kofax. [http://demo.altosoft.com/view/\(S\(mvvpjkgfvrrh3ggz3tgbh2hl\)\)/default.aspx?id=anatomicpathology&name=anatomicpathology](http://demo.altosoft.com/view/(S(mvvpjkgfvrrh3ggz3tgbh2hl))/default.aspx?id=anatomicpathology&name=anatomicpathology) (accessed June 2015)
- [4] *Anatomic Pathology Dashboard*, Viewics, Inc. <https://viewics.com/slider-item/anatomic-pathology-dashboard/> (accessed June 2015)
- [5] Chen, Y., McCormack, G., Heher, Y.K.: “Adopting the Quality Dashboard in Pathology: Real-Time Data Monitoring and Improvement”. United States and Canadian Academy of Pathology (USCAP) 105th Annual Meeting, accepted abstract and poster presentation. *Modern Pathology*, Vol. 29, pp. 488–509, <http://www.nature.com/modpathol/journal/v29/n2s/pdf/modpathol201623a.pdf> (2016)
- [6] *Clinical Pathology Dashboards*, University of Michigan. <https://www.pathology.med.umich.edu/index.php?t=page&id=1206> (accessed June 2015)
- [7] D3.js (Data-Driven Documents), <http://d3js.org/s.org/>, (accessed December 2015)
- [8] Eckerson, W.: *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*, Second Edition, Chapters 1, 6, 11. John Wiley & Sons, Hoboken, U.S. (2011)
- [9] *EORLA: About Us*. <http://www.eorla.ca/about-us/>, (accessed July 2016)
- [10] Feldt, R., Magazinius, A.: “Validity Threats in Empirical Software Engineering Research - An Initial Survey”. In: *Proc. Int. Conference on Software Engineering and Knowledge Engineering (SEKE 2010)*, pp. 374–379, KSI (2010)
- [11] Ferranti, J.M., Langman, M.K., Tanaka, D., McCall, J., Ahmad, A.: “Bridging the gap: leveraging business intelligence tools in support of patient safety and financial effectiveness”. *Journal of the American Medical Informatics Association*, Vol. 17, Issue 2, pp. 136–143 (2010)
- [12] Few, S.: *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, pp. 26 (January 2006)
- [13] Fox, B.I., Hollingsworth, J.C., Gray, M.D., Hollingsworth, M. L., Gao, J., Hansen, R.A.: “Developing an expert panel process to refine health outcome

- definitions in observational data”. *Journal of Biomedical Informatics*, Vol. 46, Issue 5, pp. 795–804 (2013)
- [14] Glaser, J., Stone, J.: “Effective use of business intelligence: leveraging your organization's business data could improve financial and operational performance--and quality of care”. *Healthcare Financial Management*, Vol. 62, Issue 2, pp. 68–73 (2008)
- [15] Golab, L.: *Sliding window query processing over data streams*. Ph.D. dissertation, University of Waterloo, Canada, 2006. <https://cs.uwaterloo.ca/~tozsu/publications/stream/cs2006-27.pdf> (accessed October 2016)
- [16] Halwani, F., Li, W.C., Banarjee, D., Lessard, L., Amyot, D., Michalowski, W., Giffen, R.: “A real-time dashboard for managing pathology processes”. *Journal of Pathology Informatics*, 7:24, Wolters Kluwer (2016)
- [17] Heatmap, *Wikipedia*, [https://en.wikipedia.org/wiki/Heat\\_map](https://en.wikipedia.org/wiki/Heat_map), (accessed July 2016)
- [18] IBM, Cognos Analytics Cognos Business Intelligence 10.2.2, *Business Intelligence Architecture and Deployment Guide 10.2.2, The Multitiered Architecture*. [https://www.ibm.com/support/knowledge-center/en/SSEP7J\\_10.2.2/com.ibm.svg.ba.cognos.crn\\_arch.10.2.2.doc/c\\_arch\\_the\\_multi-tierarchitecture.html](https://www.ibm.com/support/knowledge-center/en/SSEP7J_10.2.2/com.ibm.svg.ba.cognos.crn_arch.10.2.2.doc/c_arch_the_multi-tierarchitecture.html), (accessed August 2016)
- [19] IBM, *IBM RAVE (Rapidly Adaptive Visualisation Engine)*. [https://www-304.ibm.com/connections/blogs/basupportlink/entry/customize\\_and\\_apply\\_visualizations?lang=en\\_us](https://www-304.ibm.com/connections/blogs/basupportlink/entry/customize_and_apply_visualizations?lang=en_us), (accessed December 2015)
- [20] IBM Support, *How to improve performance by reusing cached data when running a report?* <http://www-01.ibm.com/support/docview.wss?uid=swg21344305>, (accessed August 2016)
- [21] Jarvenpaa, S.L., Dickson, G.W.: “Graphics and managerial decision making: Research-based guidelines”. *Communications of the ACM*, Vol.31, Issue 6, pp. 764–774 (1988)
- [22] Jha, A.K., Ferris, T.G., Donelan, K., DesRoches, C., Shields, A., Rosenbaum, S., Blumenthal, D.: “How common are electronic health records in the United States? A summary of the evidence”. *Health Affairs*, Vol. 25, Issue 6, pp. 496–507 (2006)
- [23] Jinpon, P., Jaroensutasinee, M., Jaroensutasinee, K.: “Business intelligence and its applications in the public healthcare system”. *Walailak Journal of Science and Technology*, Vol. 8, Issue 2, pp. 97–110 (2011)
- [24] Johnson, G.I., Clegg, C.W., and Ravden, S.J.: “Towards a practical method of user interface evaluation”. *Applied Ergonomics*, Vol. 20, Issue 4, pp. 255–260 (1999)
- [25] Karami, M., Fatehi, M., Torabi, M., Langarizadeh, M., Rahimi, A., Safdari, R.: “Enhance hospital performance from intellectual capital to business intelligence”. *Radiol Manage*, Vol. 35, Issue 6, pp. 30–35 (2013)

- [26] Kelleher, C., Wagener, T.: “Ten guidelines for effective data visualization in scientific publications”. *Environmental Modelling & Software*, Vol. 26, Issue 6, pp. 822–827 (2011)
- [27] Kimball, Ralph. The data warehouse lifecycle toolkit: expert methods for designing, developing, and deploying data warehouses. John Wiley & Sons (1998)
- [28] Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., Linkman, S.: “Systematic literature reviews in software engineering—a systematic literature review”. *Information and Software Technology*, Vol. 51, Issue 1, pp. 7–15 (2009)
- [29] Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O. P., Turner, M., Niazi, M., Linkman, S.: “Systematic literature reviews in software engineering—a tertiary study”. *Information and Software Technology*, Vol. 52, Issue 8, pp. 792–805 (2010)
- [30] Kohn, L.T., Corrigan, J.M. and Donaldson, M.S.: "Crossing the quality chasm: a new health system for the 21st century." *Washington, DC: Committee on Quality of Health Care in America, Institute of Medicine* (2001).
- [31] Kolowitz, B.J., Shresth, R.B.: “Enabling Business Intelligence, knowledge management and clinical workflow with single view”. *Issues in Information Systems*, Vol. 12, Issue 1, pp. 70–77 (2011)
- [32] Lanna, M., Amyot, D.: “Spotting the Difference”. *Software: Practice & Experience*, Vol. 41, Issue 6, pp. 607–626 (2011)
- [33] Luhn, H.P.: “A Business Intelligence System”. *IBM Journal of Research and Development*, Vol. 2, Issue 4, pp. 314–319 (1958)
- [34] McLaughlin, N., Afsar-manesh, N., Ragland, V., Buxey, F., Martin, N.A.: “Tracking and Sustaining Improvement Initiatives: Leveraging Quality Dashboards to Lead Change in a Neurosurgical Department”. *Neurosurgery*, Vol. 74, Issue 3, pp. 235–244 (2014)
- [35] Mettler, T., Rohner, P.: *Performance management in health care: the past, the present, and the future*. <https://www.alexandria.unisg.ch/214579/1/WI2009.pdf>, (accessed July 2016)
- [36] Miles, M.B., Huberman, A. M., Saldañaand, J.: *Qualitative data analysis: a methods sourcebook*. Sage (2014)
- [37] Nagy, P.G., Konewko, P., Warnock, M., Bernstein, W., Seagull, J., Xiao, Y., George, I., Park, A.: “Novel, Web-Based, Information-Exploration Approach for Improving Operating Room Logistics and System Processes”. *Surgical Innovation*, Vol. 15, Issue 1, pp. 7–16 (2008)
- [38] Nagy, P.G., Warnock, M.J., Daly, M., Toland, C., Meenan, C.D., Mezrich, R.S.: “Informatics in Radiology: Automated Web-based Graphical Dashboard for Radiology Operational Business Intelligence”. *Radiographics*, Vol. 29, Issue 7, pp. 1897–1906 (2009)

- [39] Nash, M., Pestrue, J., Geier, P., Sharp, K., Helder, A., McAlearney A.S.: “Leveraging Information Technology to Drive Improvement in Patient Satisfaction”. *Journal for Healthcare Quality*, Vol. 32, Issue 5, pp. 30–40 (2010)
- [40] Negash, S., Gray, P.: “Chapter 9: Business Intelligence”. *Handbook on Decision Support System 2*. Springer, Berlin, pp. 175–193 (2008)
- [41] OLTP vs OLAP, DataOnFocus, <http://www.dataonfocus.com/oltp-vs-olap/#oltp-vs-olap>, (accessed August 2016)
- [42] Peffers, K., Tuunanen, T., Rothenberger, M. A., Chatterjee, S. “A design science research methodology for information systems research”. *Journal of management information systems*, Vol. 24, Issue 3, pp. 45–77 (2007).
- [43] Poissant, L., Pereira, J., Tamblyn, R., Kawasumi, Y.: “The impact of electronic health records on time efficiency of physicians and nurses: a systematic review.” *Journal of the American Medical Informatics Association*, Vol. 12, Issue 5, pp. 505–516 (2005)
- [44] Power, D.J.: *A Brief History of Decision Support Systems*, Version 4.1, <http://dssresources.com/history/dsshhistory.html>, (accessed July 2016)
- [45] Rohloff, R.: “Healthcare BI: a tool for meaningful analysis”. *Healthcare Financial Management*, Vol. 65, Issue 5, pp. 100–108 (2011)
- [46] Stone, D., Jarrett, C., Woodroffe, M., and Minocha, S.: *User Interface Design and Evaluation*. Elsevier (2005)
- [47] The Data Warehouse Institute. “Module 2, Business Application Fundamentals”. *TDWI Business Intelligence Fundamentals, training materials* (2016)
- [48] Tufte, E.R.: *The Visual Display of Quantitative Information*. Graphics Pr; 2nd Edition, p. 10 (January 2001)
- [49] Vest, J.R., Gamm, L.D.: “Health information exchange: persistent challenges and new strategies”. *Journal of the American Medical Informatics Association*, Vol. 17, Issue 3, pp. 288–294 (2010)
- [50] Wadsworth, T., Graves, B., Glass, S., Harrison, A.M., Donovan, C., Proctor, A.: “Using business intelligence to improve performance”. *Healthcare Financial Management*, Vol. 63, Issue 10, pp. 68–72 (2009)
- [51] Wilkinson, L. *The grammar of graphics*, Second Edition. Springer Science & Business Media (2006)
- [52] Williams, R.: *The Non-designer’s Design Book: Design and Topographic Principles for the Visual Novice*. Pearson Education, 2014.
- [53] Zelazny, G.: *Say it with charts*. McGraw Hill (1996)
- [54] Zheeng, G., Zhang, C., Li, L.: “Bringing business intelligence to healthcare informatics curriculum: a preliminary investigation”. In *Proceedings of the 45th ACM technical symposium on Computer science education*, ACM, pp. 205–210 (2014)

# Appendix A: Contact Email Sent to Potential Participants

---

Dear member of EORLA,

We are conducting a study aiming to evaluate certain features of a business intelligence (BI) dashboard tool developed specifically to support operations managers in pathology facilities. We are looking to gather feedback on this tool from people holding managerial roles or having managerial responsibilities in pathology facilities. This feedback would be highly valuable, potentially leading to BI dashboard tools helping pathology managers to pro-actively address delays and other issues negatively impacting pathological diagnosis within their facilities. If you have managerial responsibilities within a pathology facility, please consider participating!

A business intelligence (BI) dashboard is a data visualization software application that displays the current status of an organization in terms of chosen business metrics and key performance indicators (KPIs). The BI dashboard being evaluated was designed to meet the specific requirements of pathology management, which were derived from literature and informal conversations with people holding managerial positions in a pathology facility.

The 60-minute session would include a demo of the proposed tool, a list of tasks to be accomplished with the tool, and a questionnaire to fill out. The questionnaire aims to help us evaluate the tool's usability, functionality, and potential value to practicing managers in pathology facilities. The session would take place at the University of Ottawa campus, at a time convenient for you, including evenings and week-ends. We would also provide you with a parking voucher for the duration of your participation.

If you are interested in participating, or have any questions or comments about the study, please contact Daniel Amyot.

Thank you for your consideration.

## Appendix B: Consent Letter

---

**Title of the study: Using Data Visualization and Predictive Analytics to Support the Management of a Hospital Pathology Department**

**Name of researcher:**

Wei Chen Li, graduate student, School of Electrical Engineering and Computer Science, phone 613-612-7680, e-mail: wli111@uottawa.ca

**Supervisors:**

Prof. Daniel Amyot, School of Electrical Engineering and Computer Science, phone 613-562-5800 ext. 6947, e-mail: damyot@uottawa.ca;

Prof. Wojtek Michalowski, Telfer School of Management, phone 613-562-5800 ext. 4955, e-mail: wojtek@telfer.uOttawa.ca;

Prof. Lysanne Lessard, Telfer School of Management, phone 613-562-5800 ext. 2468, email: Lysanne.Lessard@telfer.uOttawa.ca

**Invitation to Participate:** I am invited to participate in the abovementioned research study conducted by graduate student Wei Chen Li from the School of Electrical Engineering and Computer Science. Wei Chen Li is under the supervision of Prof. Daniel Amyot from the School of Electrical Engineering and Computer Science, Prof. Wojtek Michalowski from the Telfer School of Management, and Prof. Lysanne Lessard from the Telfer School of Management.

**Purpose of the Study:** The purpose of the study is to help evaluate certain features of a business intelligence (BI) dashboard tool. A BI dashboard is a data visualization software application that displays the current status of an organization in terms of chosen business metrics and key performance indicators (KPIs). The BI dashboard being evaluated was designed to meet the specific requirements of pathology management, which were derived

from literature and informal conversations with people holding managerial positions in a pathology facility.

**Participation:** My participation will consist of a single 60-minute session during which the researcher will first explain how the tool is to be used in the context of the experiment. I will then be asked to use the dashboard to answer the pre-defined questions, and finally to answer an anonymous questionnaire with 22 multiple-choice and short answer questions about my impressions regarding the experiment. The session has been scheduled for (\_\_\_place, date and time of each session\_\_\_).

**Risks:** Beyond the time needed to participate in this experiment, there are no known risks associated with my participation in this experiment greater than those I might encounter in everyday life. I have received assurance from the researcher that every effort will be made to minimize these risks by ensuring that the session last a maximum of 60-minutes and by scheduling the session at a time convenient to me.

**Benefits:** My participation in this study will provide the research with experimental data to evaluate and propose improvements to the Pathology BI (Business Intelligence) Dashboard.

**Confidentiality and anonymity:** I have received assurance from the researcher that the information I will share will remain strictly confidential by communicating the results of the experiment in an aggregated manner (e.g. “half of the participants found that the tool was easy/difficult to learn”). I understand that the contents will be used only to evaluate the Pathology BI Dashboard.

**Anonymity:** My anonymity will be protected by not recording my name or any identifiable information. If needed, data may be tagged with non-traceable numeric IDs.

**Conservation of data:** The data collected in the hard-copy questionnaire will be kept in a secure manner that will be accessed only by the researchers. The data will be kept by one of the supervisors for a period of 5 years in case of an audit.

**Voluntary Participation:** I am under no obligation to participate and if I choose to participate, I can withdraw from the study at any time and/or refuse to answer any questions, without suffering any negative consequences. If I choose to withdraw, all data gathered until the time of withdrawal will be removed.

**Acceptance:** I, (Name of participant) \_\_\_\_\_, having read and understood the above text, and having had the possibility to ask and receive complementary information on the study, agree to participate in this study.

If I have any questions about the study, I may contact the researcher or her supervisors.

If I have any questions regarding the ethical conduct of this study, I may contact the Protocol Officer for Ethics in Research, University of Ottawa, Tabaret Hall, 550 Cumberland Street, Room 154, Ottawa, ON K1N 6N5

Tel.: (613) 562-5387

Email: ethics@uottawa.ca

There are two copies of the consent form, one of which is mine to keep.

Participant's signature:            (Signature)            Date: (Date)

Researcher's signature:            (Signature)            Date: (Date)

## Appendix C: Experiment Protocol

---

*Reminder on the use of this protocol: The activities in this protocol are to be followed in the order described below for each participant. However, the time estimated for each activity may change from one participant to the other; this should be allowed and participants should not be rushed to finish any activity.*

1. Consent form (5 minutes). The participant is presented with the consent form and given the chance to ask any question or clarification before signing.
  - 1.1. If the participant declines to sign the consent form, thank him/her and give him/her a parking voucher before departing.
  - 1.2. If the participant signs the consent form, continue the experiment.
2. Demonstration (10 minutes). Demonstrate the basic functionalities of the BI dashboard tool to the participant, including those that will be required to complete the tasks listed in point 3.
3. Tasks (30 minutes). The participant is given the “Participant tasks” document (see Appendix D – Participant tasks), which asks him/her to perform a sequence of eight tasks and fill in the blanks.
  - 3.1. The researcher should verbally provide help to the participant if asked to do so, or if she observes that the participant is having difficulties completing a task. Detailed instructions for each question are written in “Appendix E – task answers and instructions” as a reminder for the researcher.
  - 3.2. The researcher should also note the duration of each task, whether help was required to complete the task, and any emerging issues that the participant may have while using the tool.
4. Questionnaire (15 minutes). The participant is given the “Questionnaire” document and asked to answer the questions provided. Remind the participant that s/he can refuse to answer any question with which s/he is uncomfortable.

- 4.1. The researcher should answer any question that the participant may have about the questionnaire or its individual questions.
5. Conclusion. The researcher collects the “Participant tasks” and “Questionnaire” documents from the participant, thanks the participant for his/her time, and gives him/her the parking voucher.

## Appendix D: Participant Tasks

---

Participant Tasks

Participant # \_\_\_\_\_

1. By checking the Pathology Dashboard, **Pending Case Analysis** component, I found that for the **Histology Lab**, the number of cases that are beyond their target processing time is \_\_\_\_\_, the number of cases that are in danger of going beyond their processing time is \_\_\_\_\_, and the number of cases that are well within their processing time is \_\_\_\_\_.
2. I want to take a closer look at the **cases that are in danger of going beyond their processing time** in **Diagnosis** for **Surgical** only. I do this by selecting the desired options from the dropdown list value prompts, narrowing the list of cases displayed. I see the list (circle one: Yes or No).
3. The list for the cases that are **in danger of going beyond their processing time** in the **Diagnosis** for **Surgical** is sorted by case created date from earliest to latest. I want to check which case has the highest Ratio, e.g., Spent hour/Target hour. By sorting descending on the Ratio column, I found that case with Accession No. \_\_\_\_\_ has the Ratio of \_\_\_\_\_.
4. I want to have a further investigation of this case and see its details. By clicking the Accession No. of this case, I see that a new case detail report is opened to me (circle one: Yes or No).
5. From the **Cases Detail List** report, I found that the number of specimens of this case is \_\_\_\_\_, and the latest event of this case is (event type) \_\_\_\_\_ at (event\_date) \_\_\_\_\_. Based on the information I got from the dashboard and the detail report, I know where the case is stuck at (circle one: Yes or No).

6. I close the **Cases Detail List** report and go back to the Pathology Dashboard. Now, I want to analysis the current pathologist workload using the **Pathologist Workload Analysis** component. From the chart, I find that Pathologist (initial name) \_\_\_\_\_ has the highest workload around (slides count) \_\_\_\_\_; among these cases, type (case type + hour range) \_\_\_\_\_ has the largest portion.
  
7. The chart in the **Pathologist Workload Analysis** component provides a direct visualization of the workload, and now I want to check the actual numbers for the workload. By changing the display type to Table, I see the table view with the workload numbers, which tells me that the Slide Count of **Surgical, < 24H** for Pathologist **P3** is \_\_\_\_\_.
  
8. Now, I want to analyze the number of cases create/finalized during the last month. From the **Case Count Analysis** component, I find that (this is an open-end question):

## Appendix E: Task Answers and Instructions

---

1. By checking the Pathology Dashboard, **Pending Case Analysis** component, I found that for the **Histology Lab**, the number of cases that are beyond their target processing time is 0, the number of cases that are in danger of going beyond their processing time is 13, and the number of cases that are well within their processing time is 269.
2. I want to take a closer look at the **cases that are in danger of going beyond their processing time** in **Diagnosis** for **Surgical** only. I do this by selecting the desired options from the dropdown list value prompts, narrowing the list of cases displayed. I see the list (circle one: **Yes** or **No**).

**Instructions:**

- From dropdown list prompt at left, select ‘**Diagnosis**’ from the options.
- From dropdown list prompt in the middle, select ‘**Amber**’ from the options.
- From dropdown list prompt at right, select ‘**Surgical**’ from the options.

3. The list for the cases that are **in danger of going beyond their processing time** in the **Diagnosis** for **Surgical** is sorted by case created date from earliest to latest. I want to check which case has the highest Ratio, e.g., Spent hour/Target hour. By sorting descending on the Ratio column, I found that case with Accession No. SP-15-003972 has the Ratio of 83.93%.

**Instructions:**

- Move cursor to the column title ‘**Ratio (S h./T h.)**’, and a pair of arrows will display.
- Click the arrows to sort the list by **Ratio**.
- If the list is sorted ascending, click the arrows again to sort by **Ratio descending**.

4. I want to have a further investigation of this case and see its details. By clicking the Accession No. of this case, I see that a new case detail report is opened to me (circle one: **Yes** or **No**).

**Instructions:**

- **Click the Accession No. SP-15-003972, and the case detail report will be automatically opened in a new window.**

5. From the **Cases Detail List** report, I found that the number of specimens of this case is   **1**  , and the latest event of this case is (event type)   **Slide Distribution**   at (event\_date)   **Jan 30, 2015 5:45:39 PM**  . Based on the information I got from the dashboard and the detail report, I know where the case is stuck at (circle one: **Yes** or **No**).

6. I close the **Cases Detail List** report and go back to the Pathology Dashboard. Now, I want to analysis the current pathologist workload using the **Pathologist Workload Analysis** component. From the chart, I find that Pathologist (initial name)   **P12**   has the highest workload around (slides count)   **80**  ; among these cases, type (case type + hour range)   **Surgical, < 24H**   has the largest portion.

7. The chart in the **Pathologist Workload Analysis** component provides a direct visualization of the workload, and now I want to check the actual numbers for the workload. By changing the display type to Table, I see the table view with the workload numbers, which tells me that the Slide Count of **Surgical, < 24H** for Pathologist **P3** is   **59**  .

**Instructions:**

- **Click anywhere in the Pathologist Workload Analysis component, and the toolbar will display at the top of the component.**
- **From the toolbar, click the arrow beside the ‘Change Display Type’ icon, and then click ‘More...’**
- **Select ‘Table’ from the list, and then click ‘OK’.**

8. Now, I want to analyze the number of cases create/finalized during the last month.  
From the **Case Count Analysis** component, I find that (this is an open-end question):

# Appendix F: Questionnaire

---

## Questionnaire

Participant # \_\_\_\_\_

This questionnaire is to be answered **anonymously**. Please, do **NOT** write in your name, but **DO** write your participant number.

All questions bellow should be answered based on the features that were discussed and/or showed during the experiment. Please, do **NOT** base your answers on previous knowledge or expected features.

For each of the questions below, *circle* the answer that best matches your opinion.

### Question 1

**Learnability** is a measure of how easy it is to learn to use a software product. Given this definition, would you say the **proposed tool** is **easy to learn**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

### Question 2

**Ease of use** is a measure of how easy it is to use a software product **after** its use has been *learned*. Given this definition, would you say the **proposed tool** is **easy to use**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

### Question 3

**Efficiency** is a measure of how quickly tasks can be performed with a software product **after** its use has been mastered. Given this definition, would you say the **proposed tool** allows you to perform tasks **more efficiently** than the tool(s) you are currently using as part of your managerial activities?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

### Question 4

**Intuitiveness** is a measure of how easy it is to understand the output or the interface of a software product. Given this definition, would you say the **proposed tool** is **intuitive**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 5**

**Tabular data** is an approach used to display summary or detail data as a list or (pivot) table. Given this definition, would you say the **proposed tool** provides useful support for **tabular data**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 6**

**Data visualization** is an approach used to display summary data as a chart or diagram. Given this definition, would you say the **proposed tool** provides useful support for **data visualization**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 7**

**Colour indicators** are a way to convey performance status through a meaningful colour scheme. Given this definition, would you say the **proposed tool** provides useful support for **colour indicators**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 8**

**Value filtering** is the ability to dynamically filter report output by user selection or input. Given this definition, would you say the **proposed tool** provides useful support for **value filtering**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 9**

**Sorting** is the ability to dynamically sort report outputs. Given this definition, would you say the **proposed tool** provides useful support for **sorting**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 10**

**Interface customization** is a way to switch between chart view and a table view, or change to different chart types according to the user's preference. Given this definition, would you say the **proposed tool** provides useful support for **interface customization**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 11**

**Real-time reporting** is the functionality supporting automatic updating of data analysis in real-time. Given this definition, would you say the **proposed tool** provides useful support for **real-time reporting**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 12**

**Drill Down** is the functionality used to explore the details related to event-level data for each case and each specimen. Given this definition, would you say the **proposed tool** provides useful support for **drilling down**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 13**

**Stakeholder-oriented views** are used to support multiple user groups and roles. Given this definition, would you say the **proposed tool** provides useful support for **stakeholder-oriented views**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 14**

**Specimen Type Level Customization** is the ability to assign different business rules per specimen type. Given this definition, would you say the **proposed tool** provides useful support for **specimen type level customization**?

Strongly Agree    Agree                      Neutral                      Disagree                      Strongly Disagree

**Question 15**

Which types of decisions, if any, do you think that the proposed tool could help you make?

**Question 16**

Would you use the proposed tool in its current state in the context of your daily professional activities?

Yes                      No

**Question 17a**

If you answered “yes” to question 16, could you briefly describe how you would use the proposed tool?



## Appendix G: Data Tier Database Structure

---

The following statements, in the Data Description Language format, formalize the database schema of the data tier discussed in Section 6.2 and can be run to create the actual structure.

```
CREATE DATABASE EORLABI AUTOMATIC STORAGE YES;
```

```
CREATE TABLE "acc_amp_event" (  
    "acc_amp_event_id" INTEGER NOT NULL,  
    "acc_id" INTEGER NOT NULL,  
    "acc_order_id" INTEGER,  
    "source_rec_id" INTEGER NOT NULL,  
    "source_rec_type" VARCHAR(255) NOT NULL,  
    "material_label" VARCHAR(255),  
    "personnel_id" INTEGER,  
    "amp_mode" VARCHAR(255),  
    "event_date" TIMESTAMP,  
    "event_location" VARCHAR(255),  
    "event_device" VARCHAR(255),  
    "event_type" VARCHAR(255),  
    "event_description" VARCHAR(255),  
    "id" INTEGER NOT NULL  
);
```

```
CREATE TABLE "acc_order" (  
    "id" INTEGER NOT NULL,  
    "acc_id" INTEGER NOT NULL,  
    "acc_specimen_id" INTEGER NOT NULL,  
    "procedure_id" INTEGER,  
    "source_rec_type" CHAR(1) NOT NULL,  
    "source_rec_id" INTEGER NOT NULL,  
    "quantity" INTEGER,  
    "billable" CHAR(1),  
    "billing_category_id" INTEGER,  
    "service_code_id" INTEGER,  
    "priority" CHAR(1),  
    "status" CHAR(1),  
    "created_date" TIMESTAMP,  
    "ordered_by_id" INTEGER,  
    "completed_by_id" INTEGER,  
    "completed_date" TIMESTAMP,  
    "billing_acct_id" INTEGER,  
    "lab_dept_id" INTEGER,  
    "procedure_methodology_id" INTEGER,  
    "instructions" VARCHAR(255),  
    "lab_note" VARCHAR(255),  
    "neg_control_id" INTEGER,  
    "pos_control_id" INTEGER,  
    "lab_dept_worklist_id" INTEGER  
);
```

```

CREATE TABLE "acc_block" (
    "id" INTEGER NOT NULL,
    "acc_id" INTEGER NOT NULL,
    "acc_specimen_id" INTEGER NOT NULL,
    "acc_order_id" INTEGER,
    "type" CHAR(1),
    "label" VARCHAR(255),
    "fixation_id" INTEGER,
    "pieces" INTEGER,
    "consult_label" VARCHAR(255),
    "acc_material_tracking_id" INTEGER,
    "print_status" CHAR(1),
    "print_dist_label_material_color_id" INTEGER,
    "embedding_instruction_id" INTEGER,
    "material_storage_id" INTEGER
);

CREATE TABLE "acc_process_step" (
    "id" INTEGER NOT NULL,
    "acc_id" INTEGER NOT NULL,
    "step_id" INTEGER NOT NULL,
    "assigned_to_id" INTEGER,
    "completed_date" TIMESTAMP
);

CREATE TABLE "acc_specimen" (
    "id" INTEGER NOT NULL,
    "acc_id" INTEGER NOT NULL,
    "container" VARCHAR(255),
    "specimen_category_id" INTEGER,
    "description" VARCHAR(255),
    "specimen_label" VARCHAR(255),
    "collection_date" TIMESTAMP,
    "recv_date" TIMESTAMP,
    "include_in_rpt" CHAR(1),
    "billable" CHAR(1),
    "topography_code_id" INTEGER,
    "coll_proc_code_id" INTEGER,
    "billing_catg_id" INTEGER,
    "consult_accession_no" VARCHAR(255),
    "status" CHAR(1),
    "icd9_topography_ind" CHAR(1),
    "acc_material_tracking_id" INTEGER,
    "fixation_id" INTEGER,
    "tmplt_profile_id" INTEGER,
    "service_code_id" INTEGER,
    "master_acc_specimen_id" INTEGER,
    "tmplt_specimen_id" INTEGER,
    "spec_rpt_tmplt_id" INTEGER,
    "illustration_section_id" INTEGER,
    "created_date" TIMESTAMP NOT NULL,
    "last_final_description" VARCHAR(255),
    "last_final_collection_date" TIMESTAMP,
    "material_status" CHAR(1),
    "material_hold_until_date" TIMESTAMP,
    "material_discarded_date" TIMESTAMP,
    "material_hold_reason" VARCHAR(255)
);

```

```

CREATE TABLE "acc_type" (
    "id" INTEGER NOT NULL,
    "name" VARCHAR(255),
    "category" CHAR(1),
    "data_tmplt_id" INTEGER,
    "dflt_rev_ctr_id" INTEGER,
    "process_id" INTEGER NOT NULL,
    "active" CHAR(1),
    "imported_case_type" CHAR(1) NOT NULL,
    "rpt_tmplt_id" INTEGER NOT NULL,
    "results_entry_method" CHAR(1),
    "type" CHAR(1),
    "dflt_specimen_profile_id" INTEGER,
    "dflt_order_profile_id" INTEGER,
    "dflt_ok_to_bill" CHAR(1) NOT NULL,
    "facility_id" INTEGER,
    "dflt_fee_sched_id" SMALLINT NOT NULL
);

```

```

CREATE TABLE "accession_2" (
    "id" INTEGER NOT NULL,
    "accession_no" VARCHAR(30) NOT NULL,
    "patient_id" INTEGER NOT NULL,
    "patient_mrn_id" INTEGER,
    "refmd_id" INTEGER,
    "delivery_loc_id" INTEGER,
    "acc_type_id" INTEGER,
    "acc_catg" CHAR(1),
    "rev_ctr_id" INTEGER,
    "NOT_USED" INTEGER,
    "billing_entity_id" INTEGER,
    "party_to_bill" CHAR(1),
    "svc_plc" CHAR(1),
    "client_billacct_id" INTEGER,
    "fee_sched_id" SMALLINT,
    "fin_class_id" INTEGER,
    "ok_to_bill" CHAR(1),
    "case_priority" CHAR(1),
    "current_status_id" INTEGER,
    "hold_status_reason_id" INTEGER,
    "status_final" CHAR(1),
    "data_tmplt_id" INTEGER,
    "rpt_tmplt_id" INTEGER NOT NULL,
    "dirty_rpt_header" CHAR(1),
    "imported_case" CHAR(1),
    "autoset_corrltn_types" CHAR(1),
    "can_load_results" CHAR(1),
    "primary_specimen_id" INTEGER,
    "results_classification_id" INTEGER,
    "primary_diag_finding" VARCHAR(255),
    "created_date" TIMESTAMP,
    "patient_age" VARCHAR(255),
    "patient_visit_id" INTEGER,
    "facility_id" INTEGER,
    "order_number" VARCHAR(255),
    "reason_for_deletion" VARCHAR(255),
    "patient_billing_account_id" INTEGER,

```

```

        "row_version" VARCHAR(255),
        "check_char" CHAR(1),
        "accession_no_sans_check_char" VARCHAR(255) NOT NULL,
        "accession_no_with_check_char_pattern" VARCHAR(255) NOT NULL,
        "hpv_results_classification_id" INTEGER,
        "rpt_related_clinical_results" CHAR(1) NOT NULL,
        "concurrent_acc_id" INTEGER,
        "added_to_panel" INTEGER NOT NULL
    );

CREATE TABLE "lab_procedure" (
    "id" INTEGER NOT NULL,
    "code" VARCHAR(255) NOT NULL,
    "description" VARCHAR(255),
    "label_description" VARCHAR(255),
    "preparation_catg" CHAR(1),
    "billable" CHAR(1),
    "service_code_id" INTEGER,
    "acc_categories" VARCHAR(255),
    "billing_catg_id" INTEGER,
    "slide_type" CHAR(1),
    "active" CHAR(1),
    "needs_pos_control" CHAR(1) NOT NULL,
    "needs_neg_control" CHAR(1) NOT NULL,
    "dflt_billing_role" INTEGER,
    "hpv" CHAR(1),
    "send_out" CHAR(1),
    "gyn_molecular_test_type_id" INTEGER
);

CREATE TABLE "tmplt_profile" (
    "id" INTEGER NOT NULL,
    "type" CHAR(1),
    "code" VARCHAR(255),
    "description" VARCHAR(255),
    "acc_categories" VARCHAR(255),
    "active" CHAR(1),
    "spec_rpt_tmplt_id" INTEGER,
    "clinical_display_option" CHAR(1),
    "clinical_display_days" INTEGER,
    "clinical_display_occurance" INTEGER
);

CREATE TABLE "process_step" (
    "id" INTEGER NOT NULL,
    "process_id" INTEGER,
    "description" VARCHAR(255),
    "type" CHAR(1),
    "stage" CHAR(1),
    "process_method" CHAR(1),
    "sort_ord" INTEGER,
    "results_interface_event" CHAR(1),
    "assigned_to_role_id" INTEGER,
    "delivery_event" CHAR(1) NOT NULL,
    "reflex_event" CHAR(1)
);

```

# Appendix H: Data Model Logical Layer Diagram

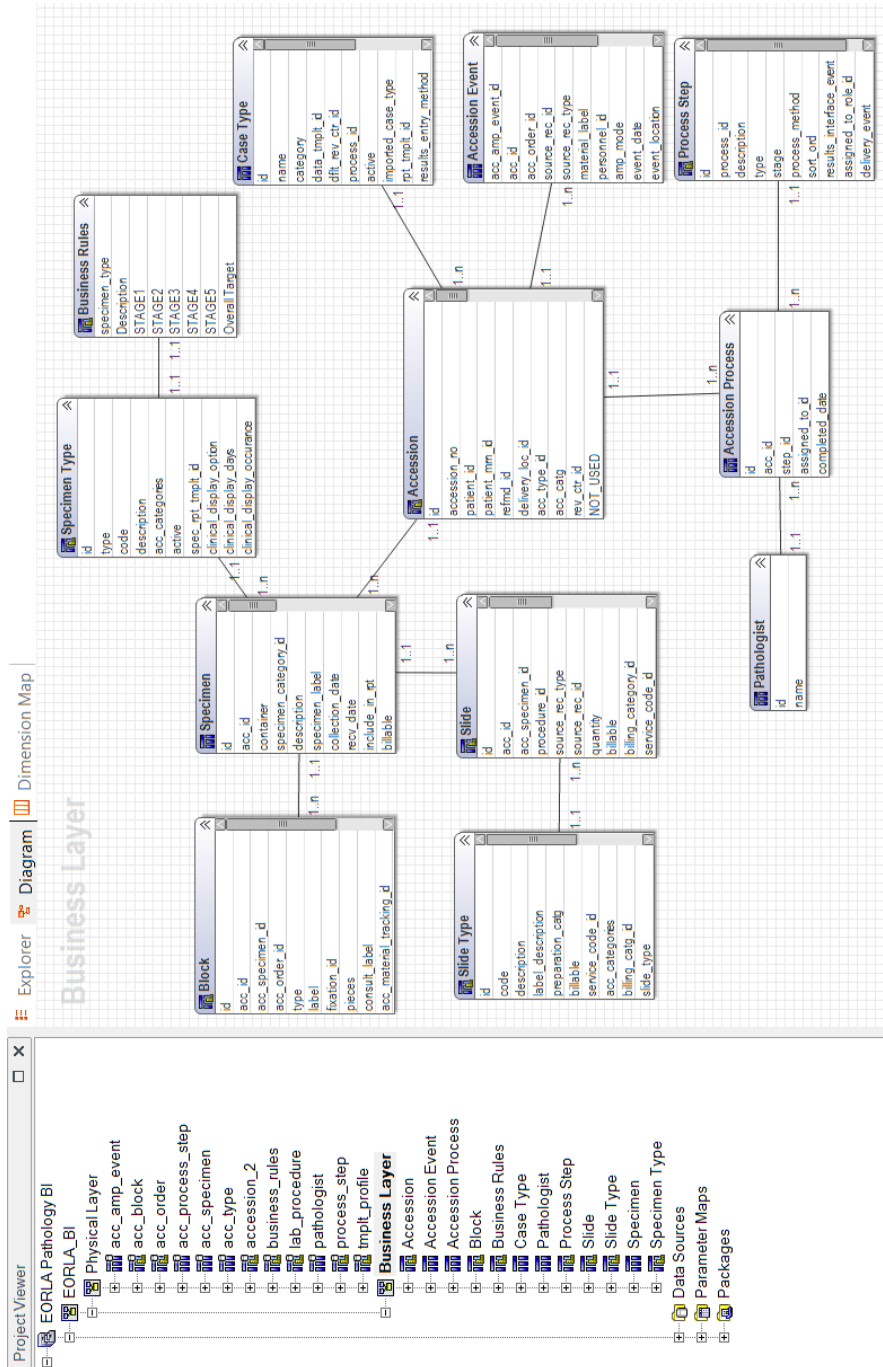


Figure 34 Data model logical layer diagram (application tier)