

# Business Intelligence in Healthcare

## Synopsis

This document outlines an approach for the adoption of Business Intelligence in the context of the Healthcare industry. Any Business Intelligence application is defined by the business questions it can answer. To that end the document advocates a “top-down” approach in gathering requirements starting from business questions through to the presentation of answers and finally the data required to support them. The document also touches on the project organization required to support Business Intelligence projects.

## Introduction

Physicians, surgeons, and managers need current, accurate and trusted information delivered in a timely manner to ensure that quality of care underlies their decision-making.

Part of the challenge in providing this information is that many healthcare workers often have many heterogeneous operational systems coexisting in their environment covering a wide and diverse range of applications from patient care, physician/provider administration, contract management for health plans through to supplies management.

The adoption and execution of a Business Intelligence (BI) strategy can fulfill a number of clear benefits:

*Consolidation and protection of data* – A single point of access to data brought together from multiple source systems can provide a single version of the truth which is one of the primary objectives of BI. Metrics calculated the same way provide the foundation of a transparent and consistent review. Workflow efficiencies can be increased through the consolidation and subsequent analysis of data brought together from clinical, administrative and financial areas. Sensitive information relating to the patient can be better protected, ensuring access only by those with the right access levels.

*Improved decision-making* – The ability to monitor the performance of doctors, departments and medical material requirements enables faster and better quality decision-making capabilities. The emphasis on analysis and accurate data brings multiple groups/individuals together and can start the push of information closer to the point of service to enhance decision-making and make data actionable.

*Effective planning* - BI can enable healthcare organizations to improve productivity with automation, cost savings, productivity and performance management. This leads to integrated planning with appropriate feedback, and in turn maximizes efficiency, effectiveness and competitiveness.

*Patient treatment* – Above all better and more accurate tracking and prediction of healthcare service utilization can improve patient care, and help optimize resource management. Increasing the potential to better maintain patient relationship management, diagnostic and treatment relationships also better facilitates timely and effective clinical decisions.

To achieve any of these goals can be a daunting undertaking and it is therefore imperative that a clear approach is adopted to ensure that the time to value is minimized and delivery of a solution is focused and achieved in small increments.

In brief the approach should be a “top-down” flow from determining who the users of the end solution will be, what questions they need to answer for their Health Care Provider and how the answers to those questions should be presented. This information can then be used to drive and scope what actual information and data is required to support those required questions and how

that data should be acquired and stored. It is then possible to accurately match the needs of the Health Care Provider with the relevant tools and technologies available that will deliver the best value.

A “top-down” approach primarily ensures that you plan to address key issues and that you can begin to form the basis of an integrated workflow moving forward. By focusing on business questions it also anticipates current and future needs of the business; being unconstrained by the detail of what specific data may or may not be easily available. By doing the reverse and working from the data upwards it is likely that at best you deliver a tactical short term solution but more likely answer random questions and almost certainly leave many unasked questions.

## **The Approach**

### **Gather Requirements**

#### **1. Questions, questions, questions**

Identify and document the relevant questions that are important for the running of their Health Care Provider or department i.e. the questions to which actual health care officers need answers to in order to monitor and manage the key metrics for their business area. For example:

- What are my nursing staff allocations and patient bed assignments?
- What is the number of patients with DVT (blood clots), infections or overstayed patients in a given unit for a given time period?

It is also important to gauge the maturity or complexity of a question to understand the likelihood of being able to answer it in the short or longer term as the BI initiative itself gains in maturity. In many cases there are more basic questions, and therefore information, that will be required to be uncovered first to provide the building blocks upon which to perform more in depth analysis. For instance the following questions imply knowledge of a significant number of more basic questions as well as the need for historical data and potentially predictive capabilities.

- What is the number of nurses and their skill levels needed in each unit based on my staff utilization, bed assignments and capacity, and projected patient flow (admissions, discharges and transfers)?
- What patients are at risk of facing harm events based on patterns or combinations of patient medical history, symptoms, diagnostics, medications, procedures and practices followed?

It is important that each phase or iteration of any BI initiative has objectives that are achievable and in line with the maturity of the Health Care Provider in relation to its knowledge and management of its data.

In some cases it will be necessary to uncover the more basic questions that go to support the recorded question. In other instances it will provide the starting point to subsequent questions that naturally follow on from the initial question. As a guide some important subject areas to tackle with healthcare business intelligence include:

- Provider care capabilities, locations, track record, cost and availability
- Encounter results, follow-ups, effectiveness, cost, time lines
- Conditions and treatment plans
- Patients, conditions, billing
- Labs and care-giving locations

#### **2. Metrics**

Define and document the metrics that are implicit or explicit within the questions documented above. These metrics are one of the primary outputs for the question gathering exercise. These metrics can then typically be grouped or aligned into the objectives of the Health Care Provider.

In the area of healthcare many of the metrics are likely to relate to counts or occurrences of something (e.g. patients, staff or conditions) particularly when looking at questions relating to the analysis of patient population trends and patterns.

### **3. Presentation**

Determine how the answers to these questions should be presented to the end user(s) and what the most appropriate form or mechanism of information consumption should be. The visualization of data is particularly important to healthcare workers who typically do not have a great deal of time for data exploration. The patterns and trends need to be instantly visible or else risk being overlooked. The second point is that most of the key decision-makers in many healthcare sectors (provider, researcher, public health authority, etc.) have been trained in some discipline which required heavy use of data. For example their background could be from a medical and/or clinical education, as an actuary or as a scientist. All of these disciplines are accustomed to data usage, and visualizing this data provides them with a powerful perspective as to what is going on.

Attempt to uncover a workflow or logical train-of-thought either from or between the answers to questions. Are there often repeated analysis paths taken by users that could help guide the analysis? For instance if an end user was looking at the results of staff utilization what are the likely next questions? Does that require context from the current data being viewed? Is it a natural 'drill-path' to a lower level of detail?

Some of the factors that should be considered are:

- Type of user
- Method of access i.e. remote or desk-based users
- Need for highly formatted, structured reports
- Level of detail
- Desired user interaction
- Need for ad hoc analysis
- Frequency and timeliness of information
- How users typically start their analysis e.g. High level and 'drill' to detail
- Distribution of information/reports

Often the best method to capture these requirements is to “mock-up” example reports in Excel or PowerPoint in order to capture this for discussion and aid visualization. It is always easier to critique something that exists rather than discuss from a blank sheet of paper.

Bear in mind the repeating cycle of Monitor, Analyze and Plan.

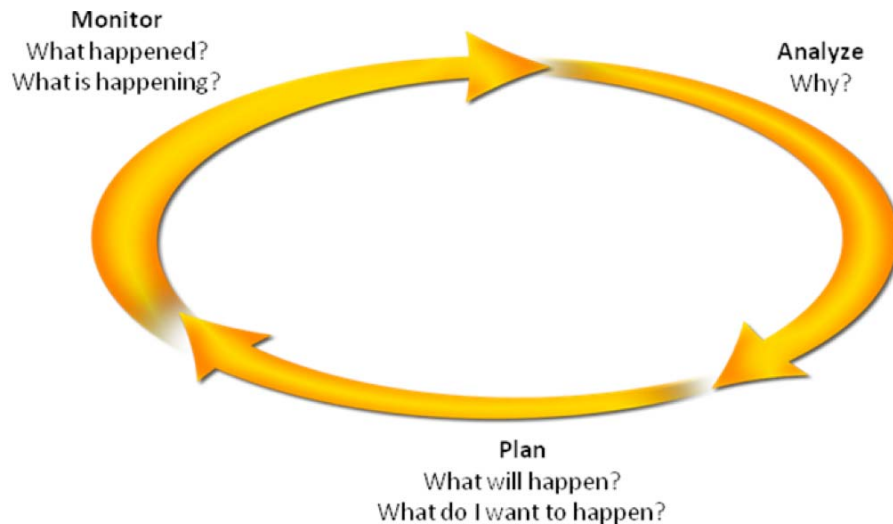


Figure 1: Performance Management Cycle

**Monitor**

This covers the use of reports, dashboards, and scorecards to allow health care employees to monitor past information. Monitor answers the “What” style of question; for example, “What happened?” “What is happening?”

**Analyze**

The next step is to use visualization tools and analytics to begin the analysis of the information behind the “What”. Analyze answers the “Why” style of questions; for example, “Why did <blank> happen?” or “Why is <blank> happening?” It is also common for users to ask better questions based on their knowledge of present business conditions.

**Plan**

Users then need to use past information and the present business conditions to make plans for the future. Plan therefore answers the “How” style of questions; for example, “How can I modify the course of action?” “How can I make the situation better?” Note that it is important that applications are integrated in order to be able to plan the future more accurately.

Another key factor when considering the presentation and consumption of data is the available technologies and capabilities. At this stage we are only looking at the top tiers of the Microsoft Business Intelligence stack i.e.

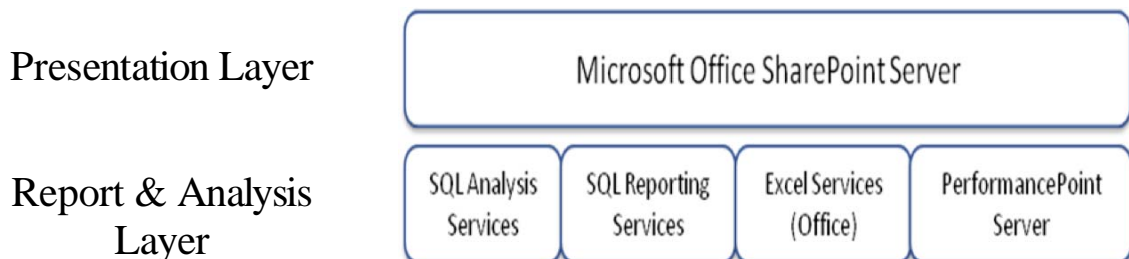


Figure 2: Microsoft BI content development environment

The consumption of information is through the Microsoft Office Suite whether that is the SharePoint portal via a browser or one of the office desktop applications.

SQL Server Analysis Services enables you to draw data from a multi-dimensional database as the source for scorecard KPIs and in-depth analysis. Access to data via Analysis Services often provides far greater performance and richer functionality than more traditional relational database access. This can be exploited through specific tools such as PerformancePoint and ProClarity but also significant advantages can be gained through using other reporting tools and even Excel on top of Analysis Services. Ad hoc and/or highly interactive reports are ideal candidates for using Analysis Services as well as providing the best performance for highly summarised data.

Structured reports (for example production or operational reporting) often associated with tactical decision-making are best produced using Microsoft SQL Server Reporting Services (SSRS). The SSRS service enables creation of traditional, paper-oriented reports and web-accessible reports which can be delivered to designated users through a self-service approach or via data driven subscriptions. In SQL Server 2005 SSRS introduces an additional tool named Report Builder which empowers less-technically inclined users to craft reports visually against a pre-defined metadata abstraction called a Report Model.

For those situations when there is a need to analyze and communicate critical business data more visually or in combination with geospatial information the capabilities of both Microsoft Office Visio® 2007 and MapPoint® 2007 can be integrated into the solution.

Microsoft Office PerformancePoint Server is a performance management application. It is an extension and evolution of earlier products Microsoft Office Business Scorecard Manager and ProClarity Analytics. It pulls together the Monitoring, Analysis and Planning elements to provide a cohesive framework for a BI application.

## Data

### 1. Data requirements

This step is focused on the analysis and documentation of what data is required to support answering the healthcare questions. This can be an involved task and may require knowledge of a variety of disparate source systems and the data held within them. It is valuable to note the complexity and rough estimated effort (hard, medium, and easy) needed to obtain this data for project planning purposes.

It is also valuable to review the key metrics derived from the captured questions and in conjunction with the increased understanding of the source system and data, validate that all the metrics necessary have been captured.

A key aspect of being able to answer any given question is the identification of all the likely dimensions that a metric will want to be viewed by (e.g. time, location, clinical service etc). It is therefore essential for the database design that the granularity of each metric/measure is known. This will drive the overall design of both the data warehouse and ETL components of the solution. It is one of the key factors that determines the data implications of the requirements and determines the most effective way to obtain the data necessary to meet those requirements.

Some of the possible dimensions within a health care organization that might be relevant are:

- Provider care capabilities
- Locations
- Track record
- Cost
- Time lines
- Conditions and treatment plans
- Patients
- Conditions

- Billing
- Labs
- Care giving locations

## 2. Database Design

Once all the requirements have been gathered and an understanding of the data required to support these requirements you need to consider how the data should be stored and aggregated.

Factors that will impact database design are the performance expectations of the end users and related to this the degree and types of user interaction. For example the provision of real-time information has design implications far beyond the creation of nightly batch reports. The final storage solution will typically be made up of a combination of relation (SQL Server) and multi-dimensional (SQL Server Analysis Services) to fulfil the expectations and requirements of the users.

## 3. Source Data

Acquiring the necessary source data for the solution involves the following steps:

- Investigate where the source data for the above is held.
- Map the fields in the source system to those required
- Evaluate gaps

Remembering the potential quantity and variety of source systems this task can be on most time-consuming and resource intensive elements.



Figure 3: Example source systems

The ability to acquire and combine this data from potentially many heterogeneous systems will affect the level of integration and therefore the depth and breadth of question able to be answered. A minimal level of integration might enable basic analysis such as:

- Number of procedures per doctor
- Hospital resource utilization
- Length of stay metrics
- Cost of care by category

A greater level of data integration may be required to address areas of analysis such as:

- Patient relationship management
- Quality outcome support
- Patient flow metrics
- Diagnostic and treatment relationships
- Optimizing duration of patient stay
- Optimizing skill mix and resource allocation to care delivery
- Ability to predict and forecast referral patterns
- Understanding cause/effect relationships i.e. Integration between patient health, recovery and follow-up after the execution of alternative treatments for similar conditions

- ## Optimizing emergency room staffing

The Extraction, Transformation and Load (ETL) tools themselves are possibly the most important of all the tools and capabilities within healthcare because as already mentioned healthcare data and data structures are typically extremely complicated. There are many standards for each subject area, be it diagnoses, procedures, treatments or outcomes. In addition, there are subject areas for which there really are no standards at all, such as measuring satisfaction, measuring accessibility to healthcare services, measuring the effects of improved outreach and so forth. It is complex and often difficult to transform the data for all of these subject areas, whether they have standards or not. To get the value out of the data requires robust and feature rich ETL tools to successfully manage this data. The capabilities within SQL Server Integration Services (SSIS) can fulfil this role.

## Project Organization

### 1. Scope

It is important that a large picture of requirements is established at the outset to determine the overall objectives and set the longer term direction. However it is also essential that only a limited number of requirements/reports are addressed at any one time i.e. think big, start small.

The following aspects can greatly affect the scope and should be considered when determining the level of work required and the deliverables:

- Number of subject areas
- Number of data sources
- Data quality
- Platforms
- Amount of customization
  - Look and feel
  - Measures/dimensions
  - Navigation/workflow
- Enterprise vs. Personal – metrics and alerts
- Security

The scope of each iteration of the project should be carefully defined and rigorously protected to avoid scope creep and delays to delivery.

### 2. Roles

The following are the typical roles required to deliver a BI project. Depending on the size of project clearly some roles could be managed by a single individual with the requisite skills.

- **Project Manager**
  - Keeping project on track, in scope, on schedule, and within budget
- **Technical Architect**
  - Overall BI system infrastructure - design, build and deploy the most appropriate technical infrastructure
- **Data Architect**
  - Overall design of the data warehouse and ETL components of the solution. Determines the data implications of the project requirements and determines the most effective way to obtain the data necessary to meet those requirements.
- **Technical Team Lead**
  - Interpreting the project's business requirements and mapping those requirements to the business intelligence application platform.

- **Business Analyst**
  - Facilitating the discovery, documentation, and prioritization of the customer's business requirements. Brings a strong understanding of the business benefits of business intelligence systems, including industry and/or functional best practices.
- **Report Developer**
  - Taking the report design specifications and implementing the universes, reports, and analytics required by the specifications.
- **ETL Developer**
  - Implementation of the translation logic outlined in the ETL design specification.
- **Test Engineer**
  - Definition and ownership of test procedures for data and report content to drive accuracy and quality of final solution.
- **Education Specialist**
  - Development and implementation of a training plan specific to the needs of the project.
  - Working with the customer to determine and deliver the most appropriate content and delivery mechanism to meet the customer's education requirements

## Summary

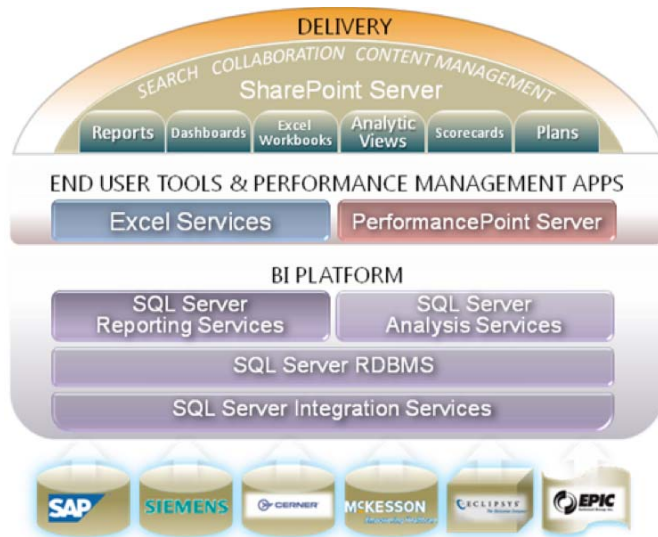
The primary purpose of any Business Intelligence initiative is to provide:

- Improved decision-making and performance by providing answers to essential questions through the use of familiar and appropriate tools.
- The ability to present the right information to the right people at the right time through well constructed dashboards and content tailored to the end users job role.
- A single version of the truth using good data management and database design built on a well integrated BI platform to provide reliable secure access.

It is worth keeping in mind that the success of any Business Intelligence initiative is defined by the business questions it can answer.

Business intelligence is no different to any potentially large or wide reaching initiative. It needs to be approached in multiple iterations allowing users to absorb information incrementally and to get involved in the project. Each iteration needs to be prioritized on those reports/questions that support the requirements being addressed at that time e.g. 10-15 high value reports. A faster implementation = faster ROI.

It is important that whilst each delivery of results may be small it is based on a solid and scalable supporting infrastructure to provide a framework for growth such as that described in the following diagram.



7

Figure 4: Microsoft Business Intelligence Platform