APPLICATION OF DATA WAREHOUSING TECHNIQUES IN THE PHARMACEUTICAL INDUSTRY

SIMON WARMAN-FREED
Niklas Data Limited
Credentials

• 12 years experience in pharma sector
• PHClinical, Data Warehousing etc.
• Combining customer needs with SAS products through our R&D programme
• Bridge between IT and Business
Issues facing Pharma Companies

• Globalisation
• Mergers / Acquisitions
• Time to market
• Harmonisation
• Accessing Data
• Accessing Information
Globalisation

- Companies are now international conglomerates
- Groups in different countries
- Sharing information
Mergers & Acquisitions

- Maximise efficiencies through effective data management
- Create common practice’s & understanding
- Increase speed of information to relevant people
- Widen access of information to relevant people
Time to Market

- Maximise sales
- Minimise use of resources:
  - people
  - time
- Transform data to information
Harmonisation

• Each study is unique which means
  – more time to produce individual programmes
  – more time taken to produce reports
  – need to spend time learning about the structure of each study
Accessing Data

- Data stored in different locations
- People are more mobile & therefore require mobility of access to data
- Data stored on different systems
- Risk
Accessing Information

• Data needs to be transformed into information
• People need access to information
• Need to share information
• Currently people rely on their data managers & statisticians to produce reports
Current state

- **Strengths**
  - No need for change of existing routines
Current state

• **Weakness**
  - Difficult to access up to date data
  - No uniform way to present data
  - No uniform cleaning of data
  - Transform data into information
  - Users need to rely on Data Managers & Statisticians to get access to data
  - Resource consuming and thereby not cost effective
Current state

• Opportunities
  – Re-use of extraction routines
  – Re-use of transformation routines
  – Re-source savings - time & costs
Current state

• Threats
  – Reports based on old or incorrect data
  – Different programmers create reports in different ways -
  – Many people “maintain” the data
Clinical Warehouse

• **Strengths**
  – Structured data
  – Safe storage of data
  – Routines for loading data warehouse
  – Controlled transformation of data
  – Can handle data from different sources
    - CDMS
    - CRO
    - Laboratory
Clinical Warehouse (Cont’d)

- **Strengths**
  - Easy access to data
  - Full Customisation
  - Applicable for all technologies
  - Same information available to everyone
  - Saves time by increasing efficiency thereby reducing overall cost
Clinical Warehouse

- **Weakness**
  - Initial Installation cost
  - Maintenance cost - reducing
Clinical Warehouse

- Opportunities
  - Reuse of extraction & transformation routines
  - Resource savings
  - Reuse of code
Clinical Warehouse

• Threats
  – Maybe the initial cost
SAS/PH-Clinical

- **Strengths**
  - Ready made reporting tool & development environment
  - Common platform to work on
  - Some Customisation possible
  - Up to date data
  - Same information available to everyone
  - Structured data
SAS/PH-Clinical

- **Weakness**
  - Need of support and training?
  - Need for templates to produce customised reports
  - Initial Installation cost
  - Maintenance cost - reducing
SAS/PH-Clinical

- **Opportunities**
  - Time savings
  - Resource savings
  - Reuse of code
  - Opportunity to receive harmonised study data
  - Change of workflow
SAS/PH-Clinical

- Threats
  - New work routines
Web application

• Strengths
  – Enterprise wide solution
  – Easy access to Intranet/Internet
  – Support global organisations
  – Easy to distribute to users
  – Easy to connect new users
  – Up to date data
  – Same information available to everyone
  – Accepted by users
  – Easy to use - web browser
Web application

• Weakness
  – New technology
  – Demand on new skills from personnel which could be difficult to find
  – Lack of pre-built functionality
Web application

- **Opportunities**
  - New work routines
    - Opportunity to most people
- No need to be at the office to run a report or view data
- By viewing data earlier also be able to respond faster to any suspicious data
Web application

- Opportunities
  - Spend time on analysing information delivered instead of just trying to get access to the information
  - Achieve higher quality and efficiency
  - New technology
    - for competitive advantage
    - for employees satisfaction
Web application

- Threats
  - New technology
  - New work routines
    - Threat to some people
Your Next Step ......

• Pharmaceutical Solutions
  - Experiences when using SAS/PH-Clinical software in clinical trials, Gunnar Magnusson, Niklas Data Sweden

• Web Enabling Ph Clinical
  - Thursday 17th June at 2.30 at the Pharmaceutical Demo Theatre

Or

• Please visit our stand
APPLICATION OF DATA WAREHOUSING TECHNIQUES IN THE PHARMACEUTICAL INDUSTRY

Simon Warman-Free - Lena Nilsson
Niklas Data Ltd, United Kingdom

ABSTRACT
In this paper we are about to give you a view of how Niklas Data are implementing different solutions either together with our customer or for our customer. The different areas we are going to cover in this paper are Clinical Data Warehouse, SAS/PH-Clinical software solutions and Web enabling of SAS/PH-Clinical software. All these approaches aim get the data to the customer faster and thereby making it more accessible to the users and by that speeding up the whole process.

When implementing these new systems and solutions for our customers we are using SAS® software modules available as well as on our skills in integrating/developing these into the new technical environment we are facing. This paper will give some more details on our experiences from building information systems and helping our customers in the pharmaceutical industry to adapt to all these new things appearing on the horizon.

NOTE
The paper presupposes that the readers of this paper are familiar with SAS/PH-Clinical software as a tool for Data Review.

For readers interested to know more about implementation of SAS/PH-Clinical software we refer to the paper "Experiences when using SAS/PH-Clinical in Clinical Trials" by our colleague Gunnar Magnusson, Niklas Data, Gothenburg, Sweden

BACKGROUND
Niklas Data has a lot of experience and engagements with customers within the pharmaceutical industry. We have been working with large pharmaceutical companies for more than a decade. During that period we have seen that this line of business has been working very much in the same way using the same tools and methodologies for a very long time. In the last year or years we have seen a change in how these companies adapt to new technologies and methodology that appears on the market.

Nearly all of the pharmaceutical companies Niklas Data have worked with the same problems occur. The main problem has been to transform and clean data into useful information and how to present information based on clinical data.

At Niklas Data, we are focusing more and more on Intranet-based systems, which will allow distributed work groups to access and share information, in particular, data coming from an existing infrastructure. Such a system could of course also be used to develop a new Clinical Data Warehouse environment. Our vision is to optimise communication of clinical research data using the power of the SAS software and to increase usage through the Intranet/Internet.

In the pharmaceutical business today we have seen that the pharmaceutical companies try to organise their clinical data more and more so that they can get access to and compare information not just within projects and studies but as well between them, independent of their global location. As a part of our long-term commitment to our customer Niklas Data are also taking part in this process using our warehouse experience and knowledge gained within other business areas. Our goal is to acts as bridge between IT and business.
GLOBALISATION

Drugs are no longer developed for one market alone but for the world market. Pharmaceutical companies are no longer national enterprises, they have become more and more international conglomerates. New markets have different demands and may need information in new ways.

As a result of this globalisation within the Pharmaceutical Industry it is becoming increasingly common for large organisations to work in distributed teams. Groups may be in different countries, or in partner companies. Key staff often work for short periods with different teams. The need to share information and work together is essential. The information systems must support the distributed nature of the work groups. This is particularly important within the Pharmaceutical Industry. To address these issues a Worldwide Information systems are needed.

TIME TO MARKET

One of the most important issues within the pharmaceutical industry today is how to shorten the lead-times to maximise the sales. One way is to speed up the process of getting data into information and to help the customers to transform data into information faster than today and by harmonisation of data be able to re-use programs between studies. Our experience is that we need to provide our customer with tools so they can achieve that and also be able to share their information. This will enable them to respond and react to the information.

HARMONISATION

A general approach to consider is to look into the area of harmonisation of study data. If the data that is used is harmonised, many of the programs made for the different solutions can be reused several times. This generates timesaving as well as resource savings.

CLINICAL DATA WAREHOUSE

A clinical data warehouse adds other demands than you have on a normal data warehouse. You may need to access data from more than one study at a time to create reports. Demands will arise to view all data within a project and to be able to do analysis on that material. To solve this issue data needs to be migrated from raw data that is in a data-entry form to a more analysis friendly format.

Below you will find a general illustration of a Clinical Data Warehouse, everything within the dotted square is the Clinical Data Warehouse.

Figure 1, Clinical Data Warehouse

EXTRACTION OF DATA

Raw data is collected through a Case Report Form (CRF) and normally entered into the inhouse Clinical Data Management System (CDMS). It’s from these CDMS that a Clinical Data warehouse normally extracts most of its data or from a similar system at a Clinical Research Organisation (CRO). In many cases laboratory data is collected in an external central laboratory and from there extracted to the Clinical Data warehouse.

Data in CDMS is normally not stored in a way suitable for analysis, it tends to be more organised in a normalised form e.g. there is more than one row per patient in the datasets. If you try to create reports based on data stored in this way you will have to invest a lot of time in transforming or merging the data. As this is done manually in a less controlled way, the risk is that the result varies from time to time.

TRANSFORMATION

After the extraction process the data needs to be transformed to format suitability for analysis. Many of these transformation programs are
written in SAS software and run after each other. The programmer must save the generated log. Even if this seems to be an easy procedure mistakes can be made. If the ordinary programmer runs these procedures the risk of mistakes may be less than if someone else has to perform the task. It may not be that obvious in which order the programs should be run to refresh the data. You can of course document in which order it should be done but the risk of running programs incorrectly is still present.

A Clinical Data Warehouse is a good way to have control over this process. Rules on how to transform the data is mainly stored in the metadata. The programs are set up to run on a schedule and in a controlled environment. In the transformation process Case Report Form (CRF) modules are grouped into analysis categories, derived variables are calculated and variables are converted to the attributes specified in metadata.

LOADING
Next step is to load the data into the analysis database. One analysis category at a time will be loaded into the analysis database. If the analysis category already exists it will be replaced by a new copy.

STORING OF THE DATA
The data is then stored in SAS datasets in the analysis database on a central NT server. Normally the data in the analysis database is organised after the clinical projects.

EXPLOIT
Easy access can be made to the analysis database and the metadata from different applications or from the development environment. How you are accessing the Clinical Data Warehouse is depending on which group you belong to. If you are a statistician, study manager or a system developer you most likely work within the development environment. If you belong to any other group you are normally working from an application.

THE DEVELOPMENT ENVIRONMENT
The main users of the development environment as mentioned before are the statisticians, study managers and system developers. The development environment contains four levels:

- Production
- Validation
- Test
- Development

All programs are written in the development environment, tested in the test level and then moved to the production environment. When the programs are moved to the production environment it’s no longer possible to edit in the program.

Why use this development environment? The purpose is to provide an easy access to the analysis database as well as enable a safe production environment.

METADATA
The metadata contains information about the data. The SAS/Warehouse Administrator is a good tool to store metadata. Normally metadata is stored for these areas:

- Which systems are delivering data to the Clinical Data Warehouse
- Information about clinical project within the Clinical Data Warehouse
- Information about studies within the Clinical Data Warehouse
- Information about the users of the Clinical Data Warehouse

EXPERIENCES FROM SAS/PH-CLINICAL SOFTWARE
Sometimes new technology gives an opportunity to change the workflow. This is what’s happening when SAS/PH-Clinical software has been used in clinical trials. Mainly there are two groups that are effected by the use of SAS/PH-Clinical software. On one side we have the Clinical Research Associates and Data Managers and on the other side the IT-staff.

Many on the clinical side are today heavily dependent on IT-staff to provide them with listings, spreadsheets or datasets with the data needed.
For the IT staff the situation is often that they do not have a common framework to develop program within. If so then it’s hard to develop general programs that can be reused.

For the two groups involved in the study there are different reasons on why a tool like SAS/PH-Clinical software can be a good solution for the problems identified above. SAS/PH-Clinical software has given the user on the clinical side the possibility to be more independent and to freely browse, search and manipulate the study data, and for the IT-staff, it is a good platform to work in, a place to create general application and many other functions.

GETTING DATA INTO INFORMATION

This had been mentioned before but nevertheless if we can provide our customer with a tool to transform data into information faster than today they could save time as well as resources. SAS/PH-Clinical software is a tool that could help us achieve these goals. It can not only provide full access to study data, but it provides the opportunity to customise the appearance of the data to fit each individual user.

By using SAS/PH-Clinical software templates the IT-staff can provide each user group with suitable tools to handle and to do analysis of their own data. This will reduce the time the IT-staff need to spend on support to users of the SAS/PH-Clinical software.

The IT staff have other demands. For them SAS/PH-Clinical software can be a great platform to use when programming. It can provide them with basic functionality like a metadata handling, an easy way of creating general applications (Templates), handling security and many other basic functions.

TIME SAVINGS

There is not going to be any immediately visible timesaving by using SAS/PH-Clinical software. Due to the fact that in the beginning time needs to be spent on learning process. The clinical staff need to learn a new way of working. They no longer need to order listings and datasets from the IT-staff, they can now access the data and produce the required listings themselves. This is a new way to work for the clinical staff and time has to be spent on training for them to make the implementation of SAS/PH-Clinical software a success.

For the IT-staff involved, the estimates are that the overall time requirement is not going to increase. The time they lose when setting up a study is gained by not having to create listings for the clinical staff. If the data structure is harmonised less time needs to be spent by each study added to the system, and programming can be kept to a minimum.

POSSIBLE RISKS

To introduce an application like SAS/PH-Clinical software with ‘change of work’ routines is always a risk. If the focus for the implementation is not to increase individual gain or at least convince them about the gain for the organisation the implementation will fail. But the success of the implementation is up to each user. The tool is there for them to use, by providing access to data or to get a platform to build upon. It’s very hard to force people into change for no apparent reason.

To reduce the risk it’s very important with information, I would say that you never can inform too much before and after an implementation phase.

EFFECTS ON STAFF

A tool like SAS/PH-Clinical software is not introduced without spending a lot of time and effort. This tool gives the clinical staff an opportunity to change the way they work. In the introduction phase there is a great need for information and support in order to help everyone to use the tool efficiently.

The effects on the IT staff are not so large in the beginning. The difference for them is that their programming effort for a study has to be done earlier than they are used to.

CONCLUSION

SAS/PH-Clinical software is very likely a tool to use in clinical trials. By using it correctly you will be able to distribute study data to anyone that needs it. The tool can be customised to suit the different users needs as well as it is an excellent platform for the IT staff to work from.
WEB ENABLING SAS/PH-CLINICAL
SOFTWARE

It is becoming increasingly common for large organisations to work in distributed teams. Groups may be in different countries, or in partner companies. The information systems must support the distributed nature of the work groups. This is particularly important within the Pharmaceutical Industry. At Niklas Data, we are focusing more and more on Intranet-based systems which will allow distributed work groups to access and share information, in particular data coming from an existing infrastructure. This Niklas Data web solution has its own metabase, designed using our experience of the clinical data process, so it could be used in systems not having SAS/PH-Clinical software.

SAS/PH-Clinical software is a very good tool for a lot of people in the Pharmaceutical Industry. But to some it will be overkill, they do not need all the features it provides. For many a web based interface would be a more suitable solution. We are here presenting a web-based interface to one of SAS Institutes products, SAS/PH-Clinical software.

ENTERPRISE WIDE SOLUTION

The demand from global pharmaceutical companies to achieve direct access to clinical data has increased a lot. For personnel that are travelling frequently there is always the problem to access the latest study data or just to access any data at all. By using an Intranet/Internet based application they only need a PC, Internet subscription and a modem to connect to “their study data”. The solution we are about to present could be the solution for people both out of the office as well as the ones that are present at the office to gain access to the latest study information.

EXPERIENCES OF INTRANET SOLUTIONS

Intranet solutions are used more and more. One reason for the wide acceptance of Intranet/Internet solutions is that today almost everyone has access to the Internet as well at offices and at home. This gives a wide acceptance among the users for the technology. It’s easy to learn and the fast access to information lets the users to respond directly to the information. It’s also providing the pharmaceutical industry with an excellent tool to share information.

ADVANTAGES USING INTRANET/INTERNET

It’s easy to distribute web-based applications to users outside the office. For the IT staff it’s easy to maintain and support, and you are getting a fast response over the network - everything is processed on the server. It’s also easy and cost effective to connect new users, all you need to get access to the study data through the Intranet/Internet is a PC, modem and an Internet subscription.

EFFECTS ON STAFF

People effected by this approach
• Investigator
• Monitor
• Medical Reviewer
• Data Manager
• IT staff
• Statistician

Historically, monitors bring piles of paper with them when visiting the investigators. Monitors out in the field will now have the possibility to access data “Up to date” for each investigator and they can easily show it to the investigator. They no longer need to bring more than the laptop with them on their travels.

The benefit for the investigator is that he/she no longer needs to wait for the monitor to arrive or the reports to be sent. The investigator can of course access their data through the Intranet themselves. This gives the monitor and the investigator time to solve more complicated issues when they meet.

Time will also be released from the Data Manager, IT staff and statisticians since the monitors, investigators and medical writers can now independently get access to and browse their own study data. They can also run/create a lot of the listings by themselves.

Anyone involved in the project does not need to wait for paper reports to arrive, the reports are available on the Intranet. Anyone that needs to see any CRF can now view registered original CRF direct, no need to visit the archive.
All involved in the project can now spend time on analysing information delivered instead of just trying to get access to the information. Last but not least important information is “Up to date” – and available when it suits you!

**ACCESS CONTROL**

Access to the system is depending on user rights defined in SAS/PH-Clinical software or if the system is used on a Clinical Data Warehouse depending on user rights defined within the systems metadata.

**HARMONISATION**

If the data that is used for the reports is harmonised the programs made for the Intranet solution can be reused several times between studies. This generates timesaving as well as resource savings.

**HOW IT WORKS**

Extractions of data can be done from PH-Clinical through an SAS/PH-Clinical software template, an extension to SAS/PH-Clinical software can be executed within the SAS/PH-Clinical software environment. This SAS/PH template will be in our scenario export metadata from SAS/PH-Clinical software to be used by the web application. Extractions can also be made from any other Clinical Data Warehouse.

All SAS Institute customers in the Pharmaceutical business do not have SAS/PH-Clinical software but they might have a Clinical Data Warehouse. By using this approach with our own metabase for the front-end, we are now able to use it with the customer and developed Clinical Data Warehouse as well.

Your local web browser sends a request for a report to a web server. The Browser request is passed forward to the SAS Application server. The request is processed by the Application Server using previous exported metadata and study data from SAS/PH-Clinical software. HTML is formatted in SAS software and sent back to the Web server. Finally the formatted HTML is sent back to your local Web browser. (See figure 2).

**CONCLUSIONS**

+ Satisfied Monitors, Investigators and Data Managers
+ High availability
+ Central storage of data
+ Reuse of programs between studies
+ Faster Validation process
- Read only
- Do not have full control of SAS/PH-Clinical software metadata

**SUMMARY**

The solutions we have been discussing are basically all addressing the same problem, how to speed up the process of transforming data to information. The Clinical Data Warehouse as well as the SAS/PH-Clinical is in many ways a tool for the information providers. Our web enabling solution is more to be addressed to the information consumer. The web solution is to be seen as a complementing tool to the Clinical Data Warehouse or to SAS/PH-Clinical in order to make data easily accessible in today’s global organisations.

Our vision is to optimise communication of clinical research data using the power of the SAS software and to increase usage through the Intranet/Internet.
ABOUT THE AUTHORS

Simon Warman-Freed
Managing Director at Niklas Data Ltd.
Niklas Data Ltd is a company within Niklas Data Group, the leading SAS Quality Partner in Europe.

Simon can be contacted via mail at:
Niklas Data Ltd.
2 Shaftesbury Court
Chalvey Park
Slough
Berkshire SL1 2ER
United Kingdom
Tel: +44-1753-732 100
Fax: +44-1753-732 110
Mail: simon.warman-freed@niklasdata.co.uk

Lena Nilsson
Pharmaceutical sector responsible at Niklas Data Ltd.
Niklas Data Ltd is a company within Niklas Data Group, the leading SAS Quality Partner in Europe.

Lena can be contacted via mail at:
Niklas Data Ltd.
2 Shaftesbury Court
Chalvey Park
Slough
Berkshire SL1 2ER
United Kingdom
Tel: +44-1753-732 100
Fax: +44-1753-732 110
Mail: lena.nilsson@niklasdata.co.uk

SAS software, SAS/PH-Clinical software and SAS/Warehouse Administrator are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. The Niklas Data logo is a registered trademark of Niklas Data Europe B.V., Amsterdam, The Netherlands.

Other brand and product names are registered trademarks or trademarks of their respective companies.