Security issues for systems used for collecting, storing and interpreting human biological data

Date received (in revised form): 4th March, 2002

David Harrison, BSc, MBA
is a Senior Consultant with ABB (formally Eutech). He has worked within the pharmaceutical and biotechnology industry in various quality assurance and computer systems validation roles. David leads ABB's 21 CFR Part 11 consultancy services and has been involved in numerous 21 CFR Part 11 projects for key blue chip pharmaceutical clients and equipment vendors.

David Harrison
ABB, Life Sciences Department,
Pavilion 12,
Belasis Hall Technology Park,
Billingham,
TS23 4YS, UK

Tel: +44 (0) 1925 741007
Fax: +44 (0) 1642 372166
E-mail: david.d.harrison@gb.abb.com

David Harrison

Abstract Over recent years there has been a substantial increase in the number of computerised systems used within the healthcare industry. The system's operation and the data produced are vital in ensuring the safety and efficacy of pharmaceutical products. At the same time the healthcare industry has been battling to keep ahead of the increasing regulatory demands of computerised system validation, one integrated component of which is system security. This paper looks at the regulatory requirements for system security and considers the impact of 21 CFR Part 11 on system security methods and practices.

Keywords: computer systems validation, physical security, logical security, biometrics, non-biometrics, 21 CFR Part 11

Introduction

The healthcare industry has seen a significant technological advance in the past ten years. The majority of pharmaceutical production-based activities are now fully automated, and the complexity and capability of analytical instrumentation have increased in line with advances in computing powers. The role and scope of computerised systems have also expanded, with a recent increase in the number of database systems whose primary role is to maintain electronic data or electronic records. From a commercial standpoint the systems range from localised computerised production equipment all the way through to global business systems. In the academic field a similar range of systems span from laboratory-based instrumentation through to Internet-based databases.

One crucial consideration that has to be addressed given our reliance on computerised systems is how to secure the system's use and maintain the integrity of the data within the system.

Regulatory requirements

For those involved in the commercial sector of the healthcare industry there are various regulatory requirements that govern the use of computerised systems. These are often referred to as GxPs (good 'x' practices, where x can stand for manufacturing, clinical, laboratory, etc.).

In the USA the main regulations are the Food and Drug Administration's (FDA) 21 CFR Parts 210[1] and 211[2] for Finished Pharmaceuticals, 21 CFR Part 820 for Medical Devices,[3] and 21 CFR Part 58 for Good Laboratory Practices for Nonclinical

Laboratory Studies.[4] In Europe the European Directives 91/EEC/44-356-412[5] is the common requirement, which is often commonly referred to as the Orange Guide.

Despite the widespread use of computer systems within the industry there is little or no detail in the GxPs on what controls and security measures are required in their use.

On a separate regulatory initiative, but one that has significantly expanded the requirements for computer systems, the FDA's 21 CFR Part 11[6] became effective in August 1997. The regulation was an enabling rule for the use of electronic signatures, but it became apparent that the scope must also be extended to put appropriate controls on the use of electronic record systems.

21 CFR Part 11 outlines the expected controls required for computerised systems that create, maintain, modify, archive or transmit electronic records, but interestingly the actual depth and detail of security controls are not specifically defined.

Guidelines and current practices

In addition to the above-mentioned regulations, industry pressure has led to the development of a number of guidelines for computerised systems. Some of the more relevant ones include GAMP 4,[7] Computerized Systems used in Clinical Trials[8] and Good Clinical Practice.[9]

Within the guidelines, the information is again inconclusive in defining what are appropriate measures for security of systems, and therefore it is usually left to the software developers or healthcare companies to decide what is appropriate. The majority of software vendors and healthcare companies have been relatively pragmatic in their approach. For example, there is no explicit requirement in the regulations for security measures to include encryption (see exclusion later for open systems), and as a result the vast majority of systems do not attempt to include complex encryption techniques. This is a practical approach and welcomed in the industry, as it is recognised that the majority of systems are used in both a closed environment and the data (although crucial to the healthcare company) is usually not at great risk from external security breaches. In the vast majority of cases the security features are required to protect the system from an internal business risk such as a disgruntled employee deleting data, or an operator fraudulently changing data to reflect the required specification.

Protection of data confidentiality

Another factor that has become increasingly important in the last few years is maintaining data confidentiality. This is particularly an issue for systems such as clinical trial database systems where the confidentiality of patient records are a key factor in averting potential business risks and contravention of data privacy legislation. It is likely that the complexity of security features put in place to protect oneself from a potential lawsuit are more stringent and a greater driving force than the requirements set forth by the healthcare regulators.

Security definitions and methods of security

System security can be categorised and defined in a number of different ways depending upon the focus of the discussion. The following sections consider the environment in which the system resides, the two different types of security within a system, and finally methods of implementing security.

Defining the systems environment – open and closed systems

Security methods should always be designed to suit the environment in which they are applied. The FDA uses the definitions of closed and open systems where security needs are viewed based upon who has potential access to the system.

A closed system is one in which the data owner has full control over who has access

to the system. Note that it defines the *data* owner (the organisation responsible for the content of the record) and not the system owner. This is a subtle difference but one that can be far reaching given the current use of external support companies and integrated systems. The majority of systems can be classified as closed systems including systems that span numerous company sites, or have external dial-in support. In these cases the organisation should still have control over who accesses the system. In any closed system the definition is only valid if there are adequate controls and policies, and that adequate measures are taken to avoid unauthorised access.

An open system is defined as a system in which the data owner does not have full control over who has access to the system. The most obvious example of an open system would one based over the Internet, where access would be through various sources such as Internet Service Providers (ISPs). In this situation any open system inherently has an increased security risk and therefore additional measures are required, encryption being a possible option.

The requirement that must be met for an open system is to ensure the authenticity, integrity and, as appropriate, the confidentiality of the data from the data's point of creation to the point of receipt.

At present open systems are less common and it is yet to be seen exactly what additional measures will be considered acceptable within the industry.

Physical security and logical security

These definitions consider access to and access within a system.

Physical security is the actual barrier that prevents an individual from accessing a system. Physical security under this definition will usually be a physical barrier such as a key-locked door or swipe card panel.

Logical security is the security that restricts an individual to certain areas within a computerised system. This will include system log-on controls such as user name and password security. Logical access inevitably relies on hidden, read-only or read-write restrictions depending upon individual-assigned access levels. For logical access to be effective there needs to be a minimum of two levels, one for the system administrator, and one for the routine user. Systems that do not have internal logical security and where all users have access to all functionality are difficult if not impossible to maintain in a secure and integral manner.

Methods of implementing security

There are essentially two main methods of access and/or authorisation security, those based around biometric controls and those based upon non-biometrics.

Biometrics involves a method of verifying an individual's identity based on measurement of the individual's physical features or repeatable actions, where those features and or actions, are both unique to that individual and measurable. The most common forms of biometrics at present are fingerprint analysers and signature dynamics. Other methods include voice recognition, facial geometry recognition and retina scans. There have been increased discussion and use of biometric retina methods for airport security in recent months.

These methods, although not currently common, will be the way of the future in the healthcare industry. The importance and potential benefits of biometrics have clearly been envisaged by the US FDA: Part 11 states that identification for electronic signatures need only rely upon a single biometric component. They have, however, stated that the biometric method must be proven to ensure that they cannot be used by anyone other than their genuine owners. This requirement is both prudent and in line with other fundamental basics of demonstrating reliable operation of computerised systems. This testing is designed to remove the concerns over systems such as the early fingerprint scanners that could be fooled by a simple photocopy of an individual's finger.

Non-biometric methods of security

encompass any other methods that are not biometric. Part 11 requires that non-biometric electronic signatures be composed of at least two distinct identification components, one of which is only executable by, and designed to be used by, the signer. The most common form of non-biometric security is the user ID and password, which is used extensively in all industries. Other variations on this theme often use a unique key card and password; the theory being that one component is physical and the other intellectual.

There are security concerns around both the unique and the confidential components of non-biometric security. User IDs are often not difficult to determine, normally they are based on employee name or number, while physical key cards although being a little more difficult to obtain, do have a tendency to remain attached to a system or 'borrowed' by other individuals.

The issues around confidential components will be apparent to all: how do you keep that password confidential, and how do you remember all of your passwords without recycling and using variations of significant or memorable personal details?

Figure 1 provides an example of a typical security operation in place for a computerised system. It shows both physical and logical security measures in place, and applies both biometric and non-biometric methods.

Criticality and risk assessment considerations

One important point that needs to be considered throughout the whole of this paper is that a system can never be made totally secure. It is not possible to *guarantee* the security of the system nor is it possible to guarantee the integrity of the data.

It is more appropriate to consider system security as one system requirement that

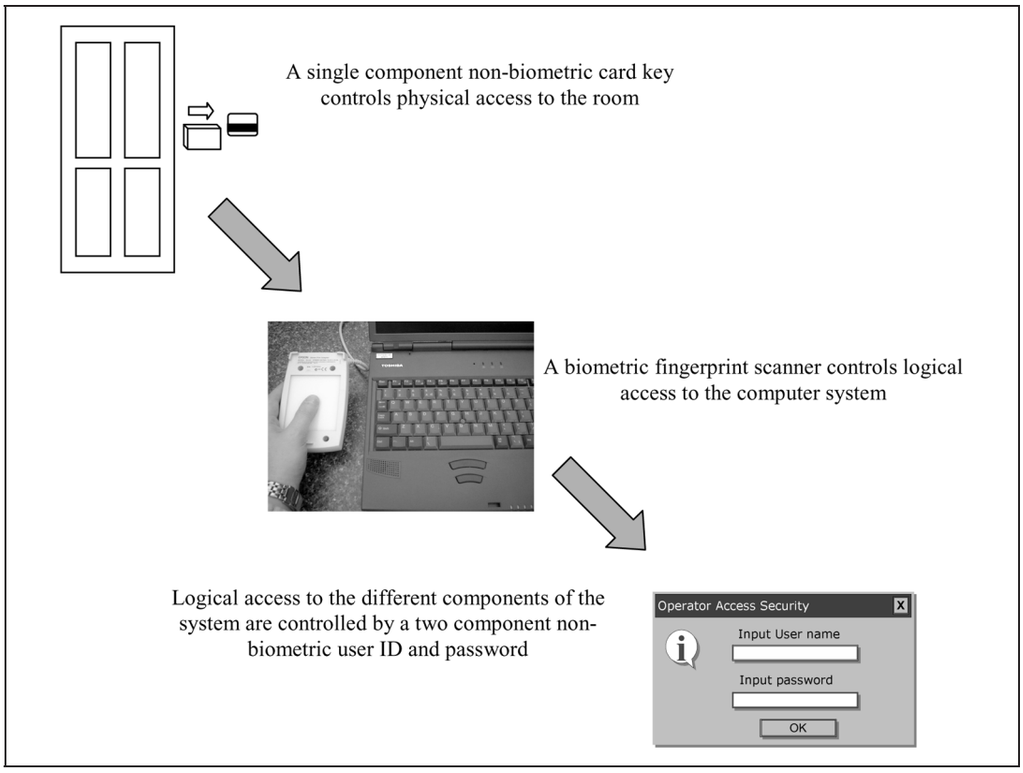

Fig. 1 Example of physical and logical security measures

must be covered by computerised system validation. It is therefore necessary to understand the goal of validation within the industry to avoid guaranteeing what cannot be delivered.

Validation is providing a *high degree of assurance* that a system purports what it is specified to do. System security is therefore providing a *high degree of assurance* that the system and data are secure from unauthorised use and that the data integrity is maintained.

Once this distinction is made it follows that all situations need to be considered on a risk analysis basis. Often systems need to be assessed and tested individually to determine what is considered an acceptable level of security given the system's current use and its current environment. This risk assessment approach is appropriate when considering both the commercial and regulatory implications of the system.

Although the commercial and regulatory implications of system security risk are inherently linked, it is useful to consider them from two separate viewpoints, business criticality and regulatory criticality.

Business criticality deals with how important the security of the data is to your organisation, what the effect will be to your financial and commercial operations if security is breached.

Regulatory criticality (often referred to as GxP criticality) deals with satisfying the regulatory requirements. The goal is to verify that the healthcare products are produced in a compliant manner and do not present any risk to public health. For computer systems this means demonstrating that the systems have been operated in a controlled and repeatable manner, and that there is accurate and integral data to verify this fact. The current regulatory requirements are based on protection of patients. This includes protection of data integrity that is used as supporting evidence for product distribution and registration, or for presenting information that is requested during an inspection.

Over recent years there have been vast improvements in the methods of security used in healthcare systems, the majority of which having been driven by vendors and healthcare companies. For example the regulations for closed systems do not actually specify the need for encryption; however, most software vendors will now include various forms of encryption to improve security. This is evidence that it is often the business criticality of a potential security breach that forces more stringent security measures. For example the financial impact of a breach of a worldwide ERP (enterprise resource planning) system is likely to be more damaging (in downtime and remediation), than the potential risk to product and patients as a results of data loss or data corruption. It should be stated, however, that the management of the organisation should take steps to ensure that data losses or corruption are detected and addressed to minimise the potential risk. Clearly this is where sensible management of risk plays a crucial role in determining a strategy for validation and ensuring security.

The conclusion here is that the business criticality drivers are usually more stringent than those of a regulatory body, and that the security measures that a healthcare company enforces upon themselves far outweigh the requirements that are asked for by the regulators.

Combining risk assessment and making efficient use of security measures

One of the dangers of having to combine both the business and regulatory risks is that the combination of the two can often result in unnecessarily strict security controls on a system. If the decision of how far the security features are taken is ultimately down to the system owner, then it is crucial to consider what is the overall goal. There are three fundamental questions that must be asked and answered.

- Do you have confidence that your system and data are secure considering the business and financial implications of a breach of security?

- Do you have confidence that your system and data are secure to meet regulatory requirements?
- Can you justify your confidence to meet regulatory requirements to an inspector?

For those organisations not subject to regulatory inspection, only the first question is relevant. It may, however, also be beneficial to consider the third question based on a purely financial and business risk audit.

In a regulated environment we need answers to all three questions. These answers need to consider both risk analysis and a considerable amount of pragmatism. It was stated earlier that it would be impossible to guarantee total security, and therefore it is safe to limit security measures to 'deter all but the most determined'.

A correctly configured and maintained firewall will prevent external access from all but the most determined, and would therefore be suitable to maintain a closed environment. In either an open or closed environment the greater risks come from internal breaches of security such as disgruntled employees. On a system using user ID and password, it is very difficult to deter people who have the time and patience to monitor colleagues logging on: eventually it is likely that the passwords will be determined. Another common risk area is access via other people's user names if they are called away or left logged on. Ideally systems should be designed to 'timeout' if the workstation is left unattended for a short period.

To combat these issues it is likely a pragmatic approach will be needed until the widespread use of biometric-based systems.

What is an inspector looking for?

The vast majority of the time inspectors will look at the suitability of the system for its task, and then whether the system is under sufficient control to satisfy that task. They will expect security to be suitable given the use of the system, its location and the number of potential users. They will begin by asking simple fundamental questions such as the following:

- Are there suitable access levels for different types of users?
- Is it possible to walk up to the system and alter its operation without detection?
- Is it possible to walk up to the system and modify or delete data without detection?
- Are there audit trails that show who created, modified or deleted electronic records and when?
- Are there obvious ways to get around the application security?

The answers to these questions are usually easily determined from initial inspection, and the security measure need to be clearly visible and apparent.

For example taking the question 'Are there obvious ways to get around the application security?', if the system is based on Windows 95 there are a number of potential problems that are likely to be apparent to an inspector:

- If data items are stored locally on the C:\ drive it is may be possible to access Windows Explorer and delete them.
- It may be possible to alter the system clock, which may then be reflected by incorrect entries in an audit trail.

It is normal in a situation such as this, where technical investigations are undertaken, that eventually an issue will be found. The deeper the investigation the more potential sources of error. This situation may lead to a regulatory non-compliance that could have been avoided by using Windows NT or saving data to a networked drive. The goal in this situation is to make the inspector feel confident in the systems security features from the initial overview, rather than feel the need to further investigate potential areas of weakness.

The effect of 21 CFR Part 11 on security features

The effect that 21 CFR Part 11 has had on the healthcare industry and software

vendors is substantial. As a result of its introduction the issues of system security have been taken more seriously. Despite the substantial impact, it is perhaps only fair to say that an initiative of this type was required to bring security concerns in the healthcare industry up to the sort of standard dictated in other IT-based businesses. The requirements and practices proposed and enforced by 21 CFR Part 11 are not new to the industry and the majority would be included in any encompassing 'good engineering practice' philosophy.

From a negative side however, the various different interpretations of 21 CFR Part 11 issues within the industry have lead to a few common misunderstandings. The outcome of the majority of these misinterpretations has been over compensating for security features rather than neglecting security.

Common 21 CFR Part 11 interpretation issues

21 CFR Part 11 deals with two separate types of systems, systems that create and maintain electronic data (records), and systems that create and maintain electronic records *and* use electronic signatures. The rule requires more stringent controls over electronic signatures than non-signature systems.

It has become apparent that many people are not adequately distinguishing between systems that hold and maintain only electronic data, and those that are used to apply electronic signatures. The inability to separate these two differing systems means that many organisations (and vendors) have applied the requirements for electronic signatures to systems that do not use signatures and only maintain electronic records. Many organisations are therefore mistakenly dismissing the suitability of their current electronic records systems as they do not meet the more demanding requirements of the signature section of the rule.

For future systems, however, it would be foolhardy to specify or design a new system that does not meet the requirements for electronic signatures, even if it does not use them. The requirements are generally considered good practice and are not technically stringent. They represent a good indication of the security features the FDA considers appropriate within the industry.

The other common mistake is that the controls specified in the Electronic Signatures section of the rule are confused with routine security features. The Electronic Signature controls are required within the system, when an electronic signature is *executed*. This is a specific activity relating to legally signing a record or data, it has nothing to do with system security (other than it must be done in a secure way).

Again many organisations are mistakenly applying the Electronic Signature requirements to their log-on security measures, when the controls are only required at the point and time of signature execution. Figure 2 shows a typical situation where this confusion may occur. The danger in Figure 2 is that users do not fully appreciate the difference between system log-on security and the legally binding signature.

A well-designed system should attempt to clearly distinguish between system security and the action of executing the signature. Figure 3 shows an ideal solution to the problem. In this situation the user should be fully aware that the fingerprint scanning action is their legally binding signature. A similar example that would be sensible and allow the distinction would be using a different user ID and password for the electronic signature execution.

Common security deficiencies highlighted by 21 CFR Part 11

There are two common system security deficiencies. The first is a lack of adequate logical security and is often only apparent for older systems. In this situation the common problem is that users are given access to functionality that they do not need access to. The ideal control philosophy would be to restrict users only to those tasks that they are authorised to perform and are

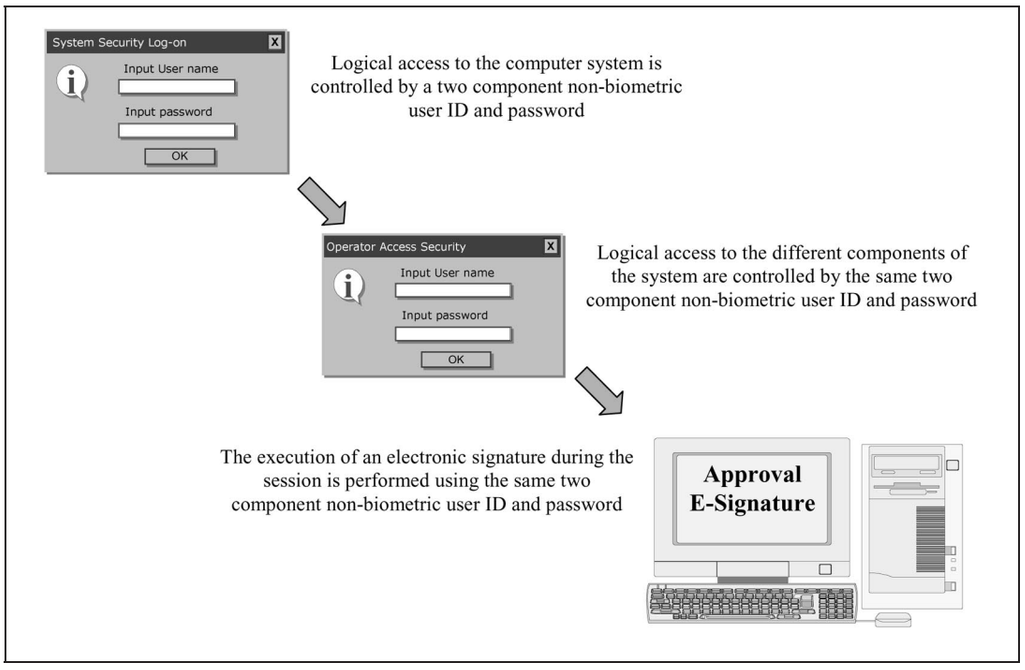

Fig. 2 Applying comparable security methods in differing situations

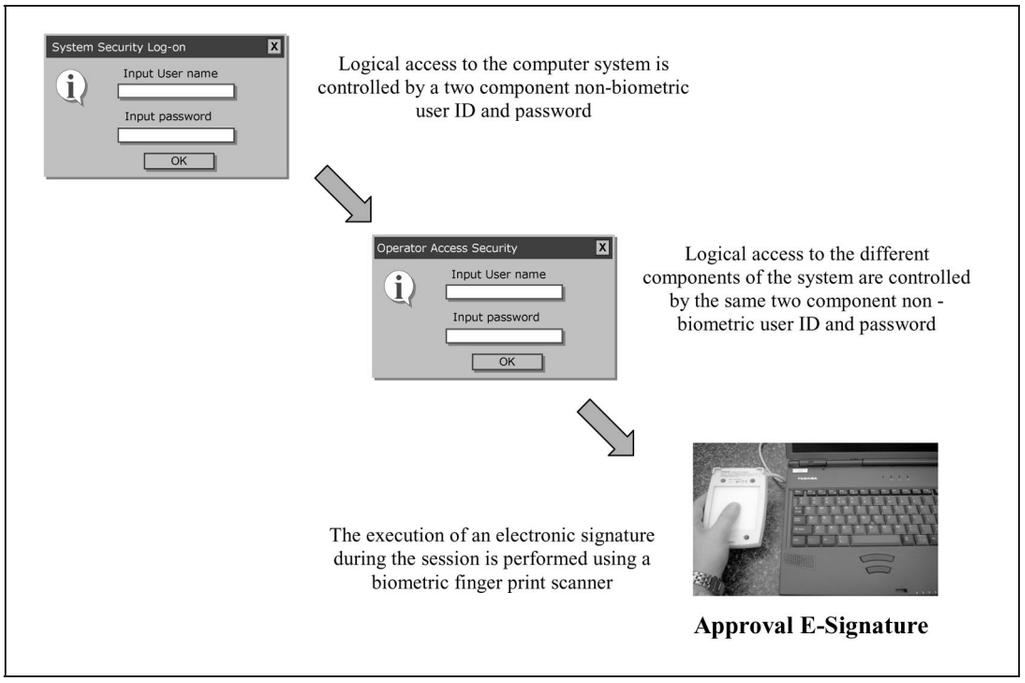

Fig. 3 Applying security methods that distinguish between system security and electronic signature execution

required to do as part of their job role. This would produce a hierarchical security structure with the system administrator having access to all functionality.

Examples of the types of activities that should be restricted to system administrators and removed for routine operators are as follows:

- Adding and deleting users from the system.
- Assigning or modifying a user's access rights.
- Access to functionality such as backing up and purging of files.
- The ability to access or configure any audit trail functionality.

This list is far from exhaustive, and is highly dependent upon the type of system being considered. Many problems are caused accidentally or maliciously by users exploring functionality that they are not trained to use, or not experienced in using.

In the majority of cases it is extremely difficult to add multilevel logical security to software unless it was designed in from the outset of the software development life cycle. This normally means that systems that provide only single level access need replacement or significant upgrade to allow secure operation.

One significant improvement in recent years has seen the introduction of fully configurable security options that meet the logical security requirements perfectly.

The second common deficiency is a lack of an audit trail. This issue is one that is increasingly becoming the focus of regulatory inspections and is a problem with many old and new systems. It could be argued that an audit trail is not actually a necessity within security features; however, the FDA has clearly stated that without a secure and fully functional electronic audit trail it is very difficult to prove the integrity of the data.

The technical requirements of designing an audit trail into a system are not actually challenging. However it is again extremely difficult to add full audit trail functionality to software unless it was designed in from the outset of the software development life cycle. Systems without an audit trail will normally need replacement or significant upgrade to allow secure and traceable operation.

An audit trail should monitor:

- when a user logs on and off;
- when a user is added to the system and when a user is removed from a system;
- when a user's access rights are changed;
- when system configurations are changed;
- when methods or processes are developed or altered;
- when methods and processes are performed;
- whenever data are manipulated or modified;
- when data are backed up, archived, purged or deleted.

It should be noted that the examples focus on system activities rather than routine operations. It is not usually critical to monitor that a user changes between screens within the software, and it may not be necessary to monitor routine operations such as changing a valve status from open to closed. These factors do not affect the integrity of the data held upon the system and are therefore unlikely to be of interest to an inspector. What is of prime importance in a system audit trail is that system security is maintained and the integrity of the data can be verified. However, when considering routine operation it is necessary to consider what is critical and what is not on a case-by-case basis and ensure appropriate actions are included in the audit trail.

It would be beneficial for software to provide a fully configurable audit trail, in which system owners can select which activities they need to review and which they can ignore. Clearly this configuration needs to be modifiable only at a system administrator level, and the audit trail will need to monitor that changes have been made to the audit trail configuration.

There have been many debates about how the audit trail should be used. At present it is usually only accessed to prove that no security breaches have occurred on the system or to identify changes to data and provide accountability for changes that have

occurred. In time it is likely that audit trails will be designed to allow beneficial and sensible filtering of entries, for example the actions of one user could be reviewed, or the activities performed on one process sample highlighted. Use of information of this type would be beneficial as supporting evidence that system operations have remained in control during use.

Procedural controls for maintaining security features

It is important to remember that security issues are not purely related to technical features of the system. The methods and procedural controls that govern the use of the system, and the way in which security measures are administered, are equally important:

- There must be procedures for setting up and deleting accounts.
- There must be procedures for ensuring that redundant users are removed from the system. This includes people leaving the company, people changing job roles, and may include when people are temporarily away from the system such as secondment or maternity leave.
- There should be controls over how the 'super user' or system administrator accounts are managed. This issue represents a considerable risk as the nature of a super user is such that they need to have access to the whole of the system and also need the functionality to administer critical functions such as system set-up and user access privileges. On most systems super users can only be controlled using independent auditing (preferably both electronic audit trails and internal quality assurance review of practices).

There are a host of other system-based procedures that help maintain the integrity and security of the data on the system. These are commonly called 'validation maintenance' procedures and comprise the following components:

- system backup;
- data archiving;
- disaster recovery;
- business continuity planning;
- change control and configuration management;
- system set-up and installation;
- system operation and maintenance.

The combination of correct technical controls and adequate procedural and operational practices result in a system that can provide a *high degree of assurance* that the system and data are secure from unauthorised use and that the data integrity is maintained.

Conclusions

21 CFR Part 11 does not actually redefine security requirements as there is still little detail on the extent of physical and logical security requirements. The biggest impact 21 CFR Part 11 has had is to force organisation to take security issues seriously and to document and control their use thoroughly.

There are a number of common misinterpretations around system security. These can usually be resolved by careful focus back to what the regulations actually require rather than what each organisation feels is suitable given their own personal business criticality risks. The security measures that a healthcare company enforces upon themselves often far outweigh the requirements that are asked for by the regulators. The danger of excessive focus on security issues is the ever-present risk of 'raising the bar', where the result is an upward spiral of the perception of what the minimum requirement is. Presently this is not something that is apparent with system security features, although it is visible elsewhere within the interpretation of 21 CFR Part 11. A pragmatic approach is required based on considered risk analysis.

Finally it needs to be recognised that successful system security requires collaboration of the healthcare manufacturers, the system vendors and the regulators. Adequate security is dependent

upon both technical controls and operational practices. The technical features are controlled by the system vendor, while the operational and procedural practices are controlled by the healthcare manufacturer. There is therefore a need for partnership to meet the system security goals and to allow compliance with the regulations.

References

1. US Regulation 21 CFR Part 210 (2001), 'cGMP in Manufacturing, Processing, Packing or Holding of Drugs; General', April.
2. US Regulation 21 CFR Part 211 (2001), 'cGMP for Finished Pharmaceuticals', April.
3. US Regulation 21 CFR Parts 808, 812 and 820 (1996), 'Medical Devices; Current Good Manufacturing Practice; Final Rule', October.
4. US Regulation CFR Part 58 (2001), 'Good Laboratory Practices for Nonclinical Laboratory Studies', April.
5. European Directives 91/EEC/44-356-412 (1997), 'Good Manufacturing Practice Rules and Guidance for Pharmaceutical Manufacturers and Distributors'.
6. 21 CFR Part 11 (1997), 'Electronic Records, Electronic Signatures. Final Rule'.
7. International Society of Pharmaceutical Engineers (2001), 'GAMP Guide for Validation of Automation Systems in Pharmaceutical Manufacture', 4th edn, ISPE, Tampa, FL.
8. FDA Guidance for Industry (1999), 'Computerized Systems Used in Clinical Trials', April.
9. Good Clinical Practice (Directive 75/318/EEC) Section 5.5 'Trial Management, Data Handling and Record Keeping'.
10. US Regulation 21 CFR Part 606 (2001), 'cGMP for Blood and Blood Components', April.