DGGF DEUTSCHE GESELLSCHAFT FÜR GUTE FORSCHUNGSPRAXIS E.V.

Digitalization of Documents in a GxP-Regulated Environment

Digitization of Documents in a GxP-Regulated Environment

1st Edition

Edited by the
German Society for Good Research Practice (DGGF)
(Deutsche Gesellschaft für Gute Forschungspraxis e.V.)

Version: 1

Date: Dezember 12th, 2011
March 1st , 2012 (English translation)

Digitization of Documents in a GxP-Regulated Environment

1st Edition

Edited by the
German Society for Good Research Practice (DGGF)
(Deutsche Gesellschaft für Gute Forschungspraxis e.V.)

Authors: Michael Bursian, Grünenthal GmbH
Thorsten Dunker, Biomedion GmbH
Britta Krusemeyer, Covance Laboratories GmbH
Klaus Lindow, Bayer Schering Pharma AG
Bernd Mohnsame, Z.A.S. Zentral Archiv Service GmbH
Bettina Quernheim, Sanofi Aventis Deutschland GmbH

Summary

As soon as you start dealing with the possibilities of digitizing documents in a GxP-regulated environment, you very quickly realize that there are no detailed guidelines for the conversion. Therefore, in a working group, consisting of representatives of archive service providers, IT service providers, contract research providers and the pharmaceutical industry, we have investigated whether the current legal situation in the various GxP sectors allow to digitize documents and then destroy the originals. In a second step, we have described the requirements imposed on the digitization process from practice. In all GxP sectors, the preparation of copies and their use instead of the originals is possible in principle. However, there must be a documented check of the result of the copying process. We suggest that this procedure should be transferred to the digitization and the degree of detail of the checking should be planned on the basis of a risk analysis. It makes sense to document the checking by means of digital signatures. The documents must be drawn up in a validated system and archived in a system operated in accordance with the requirements of 21 CFR part 11/EU GMP Annex 11 / CROMERR.

Contents

1	**Introduction**	**5**
2	**Regulatory framework**	**5**
2.1	Good Laboratory Practice	5
2.2	Good Manufacturing Practice	8
2.3	Good Clinical Practice	11
2.4	Further sources	13
2.5	Summary and interpretation of the requirements	14
3	**Organizational and technical requirements imposed on the digitizing process**	**15**
3.1	Scenarios for the digitization	15
3.1.1	Scanning for preparing a simple working copy	15
3.1.2	Preparation of a qualified copy	16
3.2	Specifying the requirements imposed on the electronic document management	16
3.3	Requirements imposed on the registry of the originals when being handed over to the digitization office	17
3.3.1	Authenticity and integrity of the originals	17
3.3.2	Requirements imposed on the list of contents / pagination	17
3.3.3	Documentation of handover / completeness / originals	17
3.4	Processing of the originals	17
3.5	Technology	19
3.6	Personnel	20
3.7	Procedure	21
3.8	Discussion of the profitability	21
3.9	File format	23
3.10	Electronic archive	24
3.10.1	Requirements imposed on the electronic archive	24
3.10.2	Archive media	24
3.10.3	Maintenance and retention of electronic data	25
3.10.4	Restoration of data after a disaster	25
3.10.5	Logical access control	25
3.10.6	SOPs	26
4	**References**	**26**
5	**Graphic representation of the scanning process**	**27**

1 Introduction

There is substantial agreement in the GxP-regulated industry that an electronic depiction of business processes is desirable. The expense for the thorough digitization of the processes that is necessary on account of the high degree of complexity has so far led to a considerable proportion of remaining media disruptions. This proportion naturally also depends on the type and size of the enterprise under consideration.

The digitization of paper documents offers an approach to providing the enterprise with at least some advantages of electronic documents. These include the quicker and simpler access to documents by authorized users and, in certain circumstances, savings resulting from discontinuation of archives for paper documents meeting statutory requirements.

However, full advantage of the merits can only be taken if the digital generated documents are equated with the original documents; in other words, the originals can be destroyed. Below we would like to give recommendations concerning the conditions required for this and outline a possible process.

This document takes into account requirements for the pharmaceutical industry as well as sponsoring contract research organizations and service providers. It can also be used for other regulated industries, although other regulations should be taken into consideration, e.g. requirements of the US EPA, CROMERR.

2 Regulatory framework

2.1 Good Laboratory Practice

The consensus document of the German Federal/State Working Good Laboratory Practice, OECD Consensus Documents No. 10 and No. 15, and the FDA Guidance for Industry Part 11 Electronic Records; Electronic Signatures — Scope and Application serve as the basis for the considerations outlined below.

All the documents mentioned above contain indications at various places that the transfer of raw data to a different medium/system is permissible, provided that the system and the transfer system have been validated and the transfer has been fully documented.

The following text passages are cited as examples:

Source	Text passage
Announcement of a consensus document of the German Federal/State Working Group on Good Laboratory Practice for the Archiving and Storage of Records and Materials **May 5, 1998** http://www.bfr.bund.de/cm/350/… …konsarch.pdf	2. Definitions and explanations … 2.1.1 Raw data …The GLP Principles permit the use of checked copies instead of original records (e.g. also reprographic reductions in size). This may be necessary in justified cases, e.g. if there is not adequate storage stability of raw data media (such as thermal printing paper). The checking of such copies for their completeness and the correct reproduction of the original raw data must be documented. In the case of paper documents, the documentation must be initialed and dated; in the case of EDP data media by means of suitable checking procedures taking the access authorizations into account. The process must be described.
	4. Documents on microfilm … After they have been filmed and their correct transfer has been confirmed, the original documents must be kept until the next inspection by the authorities, or for at least three years. Documents that are filmed later then three years after being archived can be destroyed after their correct transfer has been confirmed. In the event of an inspection by the authorities, filmed documents must be produced on request in paper form (re-enlargement). … The following are ruled out for filming: Documents that it is not possible to film satisfactorily on account of their condition or documents whose re-enlargement no longer permits complete and reliable evaluation. (Note.: Parallels to scanning could no doubt be drawn here)

Source	Text passage
Principles of Good Laboratory Practice and Compliance Monitoring **Number 10 GLP Consensus Document** **The Application of Principles of GLP to Computerized Systems (1995)** www.oecd.org	… 5. Data … Where computerized systems are used to capture, process, report or store raw data electronically, system design should always provide for the retention of full audit trails to show all changes to the data without obscuring the original data. It should be possible to associate all changes to data with the persons making those changes by use of timed and dated (electronic) signatures. Reasons for change should be given. … Where system obsolescence forces a need to transfer electronic raw data from one system to another, then the process must be well documented and its integrity verified. Where such migration is not practicable, then the raw data must be transferred to another medium and this must be verified as an exact copy prior to any destruction of the original electronic records. (Note: Here, too, again an indication that exact copies can replace the original). … 9. Archives The GLP Principles for archiving data must be applied consistently to all data types.
OECD Environment, Health and Safety Publications Series on Principles of Good Laboratory Practice and Compliance Monitoring **No. 15** **Advisory Document of the Working Group on Good Laboratory Practice** **Establishment and Control of Archives that Operate in Compliance with the Principles of GLP** www.oecd.org	… 8.2 Storage Media Records may be migrated from a computerised system onto a storage medium, e.g. magnetic tape, diskette, CD or optical disk that can be placed in a physical archive. Archive procedures should include the consideration of additional controls for the migration of electronic records from old to new media of these records. Consideration should be given to future access to the data or records stored on these media. There may be a need for special storage conditions, e.g.protection from magnetic fields.

Source	Text passage
Guidance for Industry Part 11, Electronic Records; Electronic Signatures — Scope and Application http://www.fda.gov/downloads/… …Drugs/GuidanceCompliance… …RegulatoryInformation/… …Guidances/UCM072322.pdf	… (lines 310-312) FDA does not intend to object if you decide to archive required records in electronic format to non-electronic media such as microfilm, microfiche, and paper, or to a standard electronic file format (examples of such formats include, but are not limited to, PDF, XML, or SGML).
21 CFR Part 58 Good Laboratory Practice for Nonclinical Laboratory Studies, Subpart J--Records and Reports, Sec. 58.195 Retention of records http://www.accessdata.fda.gov/… …scripts/cdrh/cfdocs/cfcfr/… …CFRSearch.cfm?fr=58.195	(g) Records required by this part may be retained either as original records or as true copies such as photocopies, microfilm, microfiche, or other accurate reproductions of the original records.

2.2 Good Manufacturing Practice

The conversion of data is not addressed in detail in national, European nor international regulatory requirements for the area of GMP. However, it is admitted at several places that true copies may be used. 21 CFR 211.180(d) is a key element for this.

According to 21CFR211.180(d) (detailed quotation: see further below), a media break is acceptable for true copies. Microfiche and microfilm are expressly given as examples for possible archiving formats. The inclusion of digitized documents that result from a scanning process is a logical continuation of these examples. This list of examples is incorporated in the GMP advisor (see also further below). Either the original or the true copy can be used for the archiving. This wording permits the conclusion that, the source material may be destroyed, when true copies are being archived instead.

However, the whole context of the quotations listed below must be taken into account. It cannot be said with absolute certainty that this interpretation is applicable to all paper documents involved, nor that the chain of argument is unequivocal.

The electronic systems that are used in a digitization process are subject to a series of requirements. These include:

- Qualified personnel with regards to both GMP and the IT systems used
- Validated procedure in implementing, using and modifying computerized systems
- Transparency and traceability with regard to who carries out what actions with electronic documents

Findings:

Source	Text passage
Arzneimittel- und Wirkstoffherstellungsverordnung (German Ordinance on the Production of Pharmaceuticals and Active Substances) (AMWHV) http://www.gesetze-im-internet.de/amwhv/	**AMWHV § 10, Para. 2,** **General Documentation** (2) If the records are made with electronic, photographic or other data processing systems, the system must be adequately validated. It must be at least ensured that the data are available for the duration of the retention period and can be made legible within a reasonable period. The stored data must be protected from loss and damage. In the event that a system is used for automatic data processing or transfer, instead of the particular responsible persons' own signature, it is adequate to state their name, provided that it is suitably ensured that only authorized persons can confirm that the particular actions have been carried out properly.
GMP AdVISOR – Reference Work for Pharmaceutical Industry and Suppliers, Status as of December 2009, Version 2.13 (Note: This work has the character of a commentary or interpretation)	**20.F.1 Documentation systems and specifications** […] Apart from the types of documents described above, there are a number of static documents, such as development reports, transfer records, validation plans and reports, training documents, etc. An archiving period must be fixed for all documents. The documents can either be archived as originals or as true copies, such as photocopies, microfiches or microfilm. […]

Source	Text passage
EudraLex - Volume 4 Good manufacturing practice (GMP) Guidelines - Annex 11 Computerized Systems http://ec.europa.eu/enterprise/… …sectors/pharmaceuticals/files/… …eudralex/vol-4/pdfs-en/anx… …11_en.pdf	1. […] Persons in responsible positions should have the appropriate training for the management and use of systems within their field of responsibility which utilises computers. […] 2. […] Validation should be considered as part of the complete lifecycle of a computer system. This cycle includes the stages of planning, specification, programming, testing, commissioning, documentation, operation, monitoring and modifying. 8. Data should only be entered or amended by persons authorised to do so. Suitable methods of deterring unauthorised entry of data include the use of keys, pass cards, personal codes and restricted access to computer terminals. There should be a defined procedure for the issue, cancellation, and alteration of authorisation to enter and amend data, including the changing of personal passwords. Consideration should be given to systems allowing for recording of attempts to access by unauthorised persons. 10. The system should record the identity of operators entering or confirming critical data. Authority to amend entered data should be restricted to nominated persons. Any alteration to an entry of critical data should be authorised and recorded with the reason for the change. Consideration should be given to building into the system the creation of a complete record of all entries and amendments (an "audit trail").
21 CFR Part 211 Current Good Manufacturing Practice for Finished Pharmaceuticals, Subpart J-Records and Reports, Sec. 211.180 General requirements http://www.accessdata.fda.gov/… …scripts/cdrh/cfdocs/cfcfr/… …CFRSearch.cfm?fr=211.180	**(**c) All records required under this part, or copies of such records, shall be readily available for authorized inspection during the retention period at the establishment where the activities described in such records occurred. These records or copies thereof shall be subject to photocopying or other means of reproduction as part of such inspection. Records that can be immediately retrieved from another location by computer or other electronic means shall be considered as meeting the requirements of this paragraph. (d) Records required under this part may be retained either as original records or as true copies such as photocopies, microfilm, microfiche, or other accurate reproductions of the original records. Where reduction techniques, such as microfilming, are used, suitable reader and photocopying equipment shall be readily available.

2.3 Good Clinical Practice

Although there have been numerous guidelines on Good Clinical Practice for a long time, the subject of scanning has not been dealt with. However, the documents "GUIDELINE FOR GOOD CLINICAL PRACTICE (CPMP/ICH/135/95)", "COMMISSION DIRECTIVE 2005/28/EC of April 08, 2005" and the FDA "Guidance for Industry Computerized Systems Used in Clinical Investigations" of May 2007 do exist, and these do indeed contain aspects of records and archiving.

The "Guidance for Industry Computerized Systems Used in Clinical Investigations" (FDA, May 2007) contains the following definition:

"A certified copy is a copy of original information that has been verified, as indicated by a dated signature, as an exact copy having all of the same attributes and information as the original."

Naturally FDA Guidance for Industry Part 11, Electronic Records; Electronic Signatures – Scope and Application is also applicable to GCP.

Source	Text passage
GUIDELINE FOR GOOD CLINICAL PRACTICE (CPMP/ICH/135/95) www.ema.europa.eu/pdfs/… …human/ich/013595en.pdf	**Original data / Source Data** **1.51 Source Data** All information from original records and certified copies of original records of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies).
	Original Documents / Source Documents **1.52 Source Documents** Original documents, data and records (e.g. hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, original recorded data from automated instruments, copies or transcriptions certified after verification as being accurate copies, microfiches, photographic negatives, microfilm or magnetic media, x-rays, subject files, and records kept at the pharmacy, at the laboratories and at medico-technical departments involved in the clinical trial).

Source	Text passage
COMMISSION DIRECTIVE 2005/28/EC http://ec.europa.eu/health/files/… …eudralex/vol-1/dir_2005_28/… …dir_2005_28_en.pdf	**Article 17** Essential documents shall be archived in a way that ensures that they are readily available, upon request, to the competent authorities.
	Article 18 Any transfer of ownership of the data or of documents shall be documented. The new owner shall assume responsibility for data retention and archiving in accordance with Article 17.
	Article 19 The sponsor shall appoint individuals within its organisation who are responsible for archives. Access to archives shall be restricted to the named individuals responsible for the archives.
	Article 20 The media used to store essential documents shall be such that those documents remain complete and legible throughout the required period of retention and can be made available to the competent authorities upon request. Any alteration to records shall be traceable.
Guidance for Industry Computerized Systems Used in Clinical Investigations **May 2007** http://www.fda.gov/downloads/… …Drugs/GuidanceCompliance… …RegulatoryInformation/… …Guidances/UCM070266.pdf	FDA is allowing original documents to be replaced by copies provided the copies are identical and have been verified as such (See, e.g., FDA Compliance Policy Guide # 7150.13).
	Definitions Certified Copy: A certified copy is a copy of original information that has been verified, as indicated by a dated signature, as an exact copy having all of the same attributes and information as the original.

Source	Text passage
s.o.	**Definitions** Original data: For the purpose of this guidance, original data are those values that represent the first recording of study data. FDA is allowing original documents and the original data recorded on those documents to be replaced by copies provided the copies are identical and have been verified as such (see FDA Compliance Policy Guide # 7150.13)[1].
	Definitions Source Documents: Original documents and records including, but not limited to, hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, copies or transcriptions certified after verification as being accurate and complete, microfiches, photographic negatives, microfilm or magnetic media, X-rays, subject files, and records kept at the pharmacy, at the laboratories, and at medico-technical departments involved in a clinical trial.

2.4 Further sources

Outside GxP, "Certified copy" is defined, for example, by the solicitor Webster Dixon LLP as follows:

"This is an official copy of an original document. A solicitor may certify the document as "a true copy of the original" only if he/she has seen and copied the original."

1 Compliance Policy Guide Sec. 130.400; Use of Microfiche and/or Microfilm for Method of Records Retention (CPG 7150.13)

2.5 Summary and interpretation of the requirements

If all the findings are taken as a basis, the following requirements for a digitization process emerge:

- Qualification or validation of hardware and software
- Comparison of original and digital copy
- Generators of digital copies identify themselves in connection with the process by means of a signature and vouch that the copy is a true reproduction of the original.

The question now arises as to how these requirements are to be interpreted. Literally, this could mean that certain pages of the documents are checked or certain pages are confirmed by the person carrying out the checking. This would be a perfect analogy to the process used in the case of paper copies. It is more appropriate to lay down the scope of the checks as a quality control in a risk analysis. Based on the type of original, e.g. the kind of paper used, the typefaces used, any pictures and graphs, etc., the likelihood of problems occurring for different types of documents should be determined in the validation process. Then the exact procedure can be laid down for the necessary quality controls regarding the various classes of documents.

In the case of simple scenarios, e.g. the digitization of normal DINA4 office paper (80g, b/w, no pictures and no special characters in the text), checking could be restricted for example to one percent of the documents if suitable hardware was used, since the risk is very low here. The other extreme case would be a scanning project, in which carbonless forms are to be digitized. With these there is an increased risk both of multiple feeding and the illegibility of the digitized copied that are obtained. In the latter case, all digitized copies must be examined individually.

Checks made in connection with the quality control must be documented and the method described in an SOP. The classification features, the scanning parameters to be used and the quality controls must be specified.

The documentation of the checks can be carried out by electronic signature on each checked document.

3 Organizational and technical requirements imposed on the digitizing process

In view of the increasing internationalization and digitization of economic life, companies have to face a constantly rising number of statutory due-diligence obligations. This applies in particular to the provision and long-term archiving of digital data in the pharmaceutical industry. In addition to the regulatory environment of Good Working Practice (GxP), the standards of FDA, EMEA and ICH may also have to be taken into consideration and the particular national provisions as for example 21 CFR Part 11, the GDPdU and the SigG as well. Due to this internationalization, it is necessary to ensure that the digitization process meets the various standards and cross-national provisions. Without professional compliance management tuned to the various requirements, it seems no longer possible to adhere to the abundance of statutory provisions, standards and procedural instructions. It is even more important for compliance management to formulate clear requirements for the data stock to be digitized perspectively ahead of a planned electronic archiving measure in order to start with the digitization process accordingly.

3.1 Scenarios for the digitization

3.1.1 Scanning for preparing a simple working copy

The objective of the simple backup is to obtain a digital working copy of the original data, in order to ensure the simple handling of information, particularly the exchange of information in the global working environment. At the same time, the original/raw data must be retained physically, centrally if possible, in compliance with GxP and in line with the statutory retention periods. An intended positive side effect of the simple working copy is the accompanying additional protection of the original documents with regard to completeness, intactness, security against manipulation etc. This protection is achieved by no more need to move originals out of the archive. The digital working copy can be provided instead.

The simple working copy does not impose any specific requirements on the digitization process, since the original of the scanned documents must be kept so that reference could be made to them in case of doubt. A completeness check and pagination of the documents, empty reverse sides of pages, an electronic signature and especially a detailed description of the digitization and work processes could therefore be dispensed with. Simple digitization with a resolution of 200 - 300 dpi black/white in PDF format can therefore be regarded as adequate. In order to ensure a sufficient accuracy and speed of access, a clear indexing specification must be drawn up before the digitization work begins.

3.1.2 Preparation of a qualified copy

The purpose of preparing a qualified copy is to be able to provide a duplicate that reflects the contents of the original document 1:1 in documentable form.

The qualified copy differs from the simple copy in that way that completeness of the digitized copy has to be guaranteed and documented. Therefore, prior to digitization, all the originals must be checked for this. A list must be drawn up that shows which documents have been collected in a file and how many pages are involved in each case. Each page must be paginated. The scan operator confirms the correct digitization of the stock of documents to the sponsor by means of his handwritten signature on the list of contents. In addition, the checking procedure is documented, as described in 2.5, by means of a signature on all the documents that have been checked.

The requirement formulated in this way leads to the following recommendations concerning the digitization processing the case of average documents: resolution of 300 dpi, color, PDF/A format, duplex (scanning of the empty reverse pages as well including an electronic signature).

3.2 Specifying the requirements imposed on the electronic document management

Coordination between technical department and archivist regarding:

- Keywording
- Fixing the retention period (at least statutory retention period)
- Fixing the access rights
- Decision on deletion of the documents

3.3 Requirements imposed on the registry of the originals when being handed over to the digitization office

3.3.1 Authenticity and integrity of the originals

Certain requirements must already be fulfilled at the time when the documents are handed over to the archives. For example the authenticity and integrity of the originals can only be confirmed by the department/company handing them over. The latter, too, only the department handing over documents can confirm whether the documents are in the correct order and complete. This is why the handover is governed by the following minimum conditions:

- List of contents
- Pagination
- Documentation of handover
- Confirmation of completeness
- Originals

3.3.2 Requirements imposed on the list of contents / pagination

In order to enable the archivist or the digitization office to ensure that the documents are complete, the originals must be paginated. There are two ways of doing this. The originals can be paginated continuously from the first page to the last and a list of contents is provided accordingly.

If the originals are already numbered from beginning to end, this can naturally also be reflected in the list of contents. For example: Analysis xy 23 pages, measuring instrument printouts Instrument yz 27 pages, etc.

Most often it makes sense to draw up a document template for a list of contents.

3.3.3 Documentation of handover / completeness / originals

When the documents are handed over to the digitization office, it must be confirmed in writing that the documents are complete. It is confirmed at the same time that all original documents or copies confirmed in them have been handed over. When the documents are handed over, the digitization office checks the documents against the list of contents and confirms the receipt. At this point, the responsibility for the further secure administration of the documents received passes to the digitization office.

3.4 Processing of the originals

After the requirements imposed on the digitized data and their registration have been clarified, it will become obvious in the course of the analysis of the stock of originals that the originals in a GxP-regulated environment are usually very heterogeneous.

Examples are given below of some of the possible peculiarities of the originals:

- Pages secured with staples or paper clips
- Documents joined with eyelets or sealed

- Documents in envelopes
- Vouchers or documents in plastic pockets or sheets
- Carbonless paper, possibly also colored
- Torn originals possibly with partial loss of information
- Originals attached or stuck
- Originals with writing on the front and / or reverse, possibly in color
- Distorted or poor print

Agreement must be reached with the competent department, if the following circumstances arise:

- Originals with pages that are not, or incorrectly, paginated
- Originals that have obviously been filed in the wrong order
- Originals that are obviously missing

The great number of different types of originals and formats means that the originals must be analyzed exactly in order to ensure that any loss of data can be ruled out in the digitization process. Particularly when digitizing existing archives, it is advisable to preview the archived records prior to document processing . For example, there may be originals that cannot be digitized because some data have been lost or their consistency does not permit to be reproduced completely and reliably . Documents of this nature must continue to be archived in paper form.

Apart from originals that are available in physical form, there is also a considerable amount of information that is already available in digital form, their file format also have to be analyzed with a view to electronic archiving.

The following are the minimum requirements that have to be fulfilled in connection with the processing of originals.

- Removal of metal: Foreign bodies must be removed from each original (staples and paper clips, sticky strips, sheets, Post-its, etc.). The order of the originals must be retained. The originals must be removed from their files or containers such as folders and boxes.
- Smoothing: If necessary, the originals must be smoothed; tears must be repaired as far as possible using special adhesive.
- Particularly thin paper must be placed in a special scanner sheet in order to prevent the original from damage.
- Stacking: The originals must be put together in piles of documents. The pagination and sequence of the originals must be checked and, if necessary, corrected.
- Originals that are difficult to read or originals where information has obviously been lost, for example due to tears, are supplemented by the insertion of appropriate DIN forms.
- The originals must be uniformly arranged in relation to the edge of the scanner.

- Pre-sorting: Where the originals are heterogeneous, it can be assumed that various scanners will have to be used for the digitization of a document. The originals must therefore be sorted within the pile of documents. Originals of a document that cannot be scanned on the main scanner but for example on a continuous or large- format scanner will be removed from the pile of documents. A dummy original will be inserted in its place. The dummy will receive a clear identification assigning it to the original. The digitization copies generated on various scanners will be combined again to a single document by EDP using this identification. The dummies themselves will be deleted again after successful combination.
- Re-sorting: After being scanned, the originals are re-sorted back 1:1 to their containers. This is necessary, since, on the one hand, the originals must be kept until the next official inspection, and for at least three years, after the digitization or filming and confirmation of the correct migration; on the other hand, in case there are complaints about the digitized copies in the course of a quality check rescans might be necessary.

3.5 Technology

The scanning technology to be used depends on the condition, format and volume of the documents to be processed. Priority must be given to the quality of the scanning results in relation to an increased throughput. This would be achieved by means of the appropriate design of the mechanics of higher quality scanners, ultrasonic sensors to prevent double drawing-in and a paper guide to avoid damage to the original.

The heterogeneity of the documents can lead to a necessity of different scanner models at the same time, for example:

- Large-format scanner up to DIN A0+, continuous, color possibly with a transparency unit, min. resolution 600 dpi
- High-performance scanner, duplex, color, with ultrasonic sensors, with automatic stop function, min. resolution 600 dpi, for normal A4 documents
- Continuous scanner DIN A3, color, min. resolution 300 dpi, for example for ECG strips
- Flatbed scanner DIN A3+, color, min. resolution 300 dpi, for bound documents
- Additionally possible:
    - Microfilm-/fiche scanner
    - X-ray film scanner

Just as important as high-quality hardware is the corresponding software.

There are already a number of scanning software providers that combine a large proportion of the software components listed below in a single product. Decisions must no doubt be made individually by the particular company as to what hardware

and software products lead to optimum results in use with other existing applications and the IT infrastructure.

The following software components are suggested in the framework of the digitization process recommended by the authors:

- Scanning software with individual control of the original
- Full text recognition
- Barcode recognition
- Electronic signature
- Date/time stamp
- Database
- PDF converter
- Indexing software

3.6 Personnel

The personnel to be employed for the digitization must be trained regarding the regulatory requirements and the process used.

Precisely the apparently simple activities like preparing the originals in a GxP-regulated environment necessitate the employment of appropriately qualified personnel and cannot as a rule be carried out by temporary staff, if the aimed result is high quality.

In addition fundamental archivist's abilities/skills, the personnel employed must also have a sound knowledge of a registration, the digitization process per se and conservational skills.

When digitization projects are awarded to external service providers, their personnel must be obligated to data confidentiality according to Section 5 of the German Federal Data Protection Act and bound to secrecy under Section 203 of the German Penal Code. Furthermore the personnel must receive instruction in the regulatory framework according to Section 2. Maybe a clearance certificate will also have to be provided.

3.7 Procedure

At this point, it is not possible to provide a description of a generally binding procedure. Instead, the procedure in question must take into account the individual needs of each company. Attention will merely be drawn to some important aspects that are generally valid.

The procedure must ensure that any loss of information during the transfer is ruled out.

Therefore adequate controls and quality measuring points must be implemented. A 100% control of the scanning result is necessary. For this purpose, each page must be compared with the pertinent original. The 4-eye principle must be ensured in the GMP and GCP sectors; a simple control is adequate in the GLP sector. All the controls must be documented by means of an advanced electronic signature.

The procedure must be described and documented as such in detail.

The scanning process, the subsequent data processing and the long-term archiving must be carried out in a GxP-compliant environment. This involves:

- Validation of hardware and software including change control procedures
- Access control at logical and physical levels
- Backup and disaster recovery procedures
- Virus protection and firewall
- Training of all administrators and users of the system
- Audit trail

3.8 Discussion of the profitability

Let us recall the following question at this juncture: “What benefits do we expect at all from electronic archiving?”

There is in fact only one answer regarding the ultimate result – at least from the angle of controlling: cost savings.

Is that in fact the case? And how is this to be achieved?

Essentially, the supposed, potential cost-saving effects focus on three sectors:

1. Faster availability of required information at all times by means of improved access and search possibilities and worldwide access
2. Space/room savings for paper archives
3. Greater security by means of several, decentralized held backups

If the life cycle of a document is considered, here for example an SOP, it is accessed most frequently during the time that it is being drawn up to the point when it comes into force. After it has been implemented, it can be assumed that it will be accessed considerably less frequently for the duration of its validity – with the exception of any revisions – up to the point where it is retired. For the duration of the following retention period, which is as a rule the longest part of

the life cycle, it can be assumed that it will be accessed very much less frequently.

The situation is similar in the case of the life cycle of most other documents in the GxP- regulated environment. It can be concluded from this that the advantage of rapid availability would in fact only be beneficial for a very limited period of time. However, since SOPs and most other documents are nowadays drawn up in digital form rapid, digital availability is given anyway. Therefore, the supposed cost advantage is at least partially obsolete.

The supposed cost savings with regard to space/room for the originals must be balanced against any digitization costs, but at least storage costs, migration and conversion costs, the administrative expense, the investment and operating costs, taking into consideration the lifespan of the electronic storage media and the qualification and validation costs.

The argument in favor of the greater physical security resulting from the decentralized storage of secured media is only valid if it is ensured that, in cases of doubt, the electronic storage medium used is still legible/readable on the long-term. On the other hand, the modern design of archives, particularly of archives of GxP-specialized archival services, normally have automatic firefighting facilities based on gas and other precautionary technical measures that virtually rule out the total loss of original documents.

Therefore to what extent electronic archiving is to be preferred to paper-based archiving from an economic point of view can (in the final analysis) only be decided with reference to a specific company. An alternative to be considered in economic terms is undoubtedly the physical archiving of old files at a GxP-specialized archive service provider with a “scan on demand” service.

External service providers consider information management with a view from outside and can in this way offer impartial solutions for all documents depending on requirements – no matter whether they are based on paper or are digital. Sensible digitization of old files according to requirements (scan on demand) paired with archive storage capacities that increase as required and storage space in association with certified safety standards mean that external service providers seem to be the right choice for many pharmaceutical companies. However, it is clear that initially the companies themselves must lay down their strategy on information management by making management decisions and include them in their long-term corporate strategy.

3.9 File format

In the past, the TIFF G4 and JPEG formats have been used frequently for digitizing paper documents. These file formats can continue to be used, although they have disadvantages in complex projects:

- Neither of the two formats has a uniform standard for the administration of metadata.
- Digital signatures cannot be administered in the file in either file format; they have to be stored in a separate file.
- The full text cannot be stored in the file in either file format.
- There are no uniform standards for the administration of multipage scans.
- Depending on the scanned document, the file formats have various advantages and drawbacks. For example TIFF G4 can be readily used in the case of black/white originals; in the case of colored originals, JPEG is preferable on account of the higher compression rates. Therefore, when projects are complex, there is frequently confusion between file formats with corresponding problems for the user, since viewers have to be kept that reproduce all file formats correctly.

The problems associated with these file formats are solved by the ISO Standard PDF/A (ISO 19005-1: Document Management – Electronic document file format for long term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)). PDF/A is a file format that supports high compression rates with high picture quality for black/white and colored document originals. When OCR technologies (text recognition) are used, the recognized full text can be stored in the file; the file can then be searched for terms with the aid of the usual standard software.

PDF/A is not an independent specification in the narrower sense, but a recommendation for the use of the PDF 1.4 Standard. Some aspects of the standard are recommended (for example bookmarks and full text); others should not be used in favor of the long-term stability of the file format (for example multimedial contents, JavaScript, …). PDF/A exists in two variants that are built on one another, PDF/A-1a and PDF/A-1b.

Essentially, PDF/A-1b requires the exact reproduction of the file contents; PDF/A-1a imposes additional requirements on the navigation possibilities in the document (list of contents with the possibility of navigation, bookmarks, …). When documents are being digitized, it will not as a rule be possible on economic grounds to generate PDF/A-1a, since the necessary metadata are not available for the digitization. However, for conversion of electronic documents, for example from Microsoft Word, the adherence to PDF/A-1a is possible at low expense, if Microsoft Word for example was used correctly for setting up lists of contents. The latter can be taken over completely when establishing PDF/A-1a and simplify navigation in long documents.

When using full text recognition, care must be taken that the recognized full text does not cover the digitized document, but is concealed in the file and used only for search purposes. Otherwise, it is possible that wrongly recognized full text information will distort the scan.

3.10 Electronic archive

3.10.1 Requirements imposed on the electronic archive

Electronic documents can of course only be trusted if they are available unchanged in the long term. Apart from the file format discussed in 3.9, the following aspects should be borne in mind:

- Protection of the data from changes by means of suitable media or software.
- Long-term stability of the archive media used. All media are subject to aging and the data must be copied from time to time.
- Logical access protection from unauthorized reading access to the archive
- Physical access protection of the facilities of the electronic archive
- Measures against disaster, as for example fire in the archive facilities
- Evidence of completeness of the data in the archive
- Traceability of the data in the archive
- Computer system validation of the system used
- Written procedural instructions for the use of the system

3.10.2 Archive media

There are a number of media that are suitable for electronic archiving. In practice, particular use is currently being made of optical media (UDO[2]/UDO2 media should be mentioned here) and solutions based on hard discs. A condition for their use in connection with electronic archiving is that the media are protected against deliberate and accidental changes. In the case of optical media, this is an attribute of the medium (TrueWORM)[3]; in the case of hard discs, suitable software must protect the data from changes (SoftWORM)[4]. The use of a simple combination of hard discs (RAID) is not advisable without further measures, since, in the first place, this technology does not have any mechanisms for the protection of data from changes by authorized users (for example system administrators). In this case, the use of a hard disc-based

[2] UDO = Ultra Density Optical, an optical medium for archiving data, comparable to the DVD
[3] TrueWORM = True Write Once Read Many, denotes the property of media that cannot be changed on account of physical characteristics, for example CD-R or DVD-R
[4] SoftWORM = Soft Write Once Read Many, denotes a process in which physically changeable media, for example hard discs, can be protected from changes by means of software. For example electronic signatures and time stamps are used for this purpose.

long-term archiving solution of a specialized manufacturer[5] or the use of standard hardware with additional software for long-term archiving purposes[6] is advisable.

Both optical media and hard disc-based systems can be used here with regard to regulatory aspects.

3.10.3 Maintenance and retention of electronic data

All archive media are subject to aging/deterioration. The media used must be regularly examined for wear and tear and exchanged in good time before the expiry of the storability specified by the manufacturer. It is important to meet the requirements regarding storage temperature and atmospheric humidity specified by the manufacturer, in order to achieve the specified storability of the media. The process of data migration to new media must be validated.

3.10.4 Restoration of data after a disaster

Systems for electronic archiving can be destroyed by fire, burglary or similar occurrences. It is therefore important to be prepared for such an occurrence. When optical media are being used, it is desirable to prepare at least two copies in each case and to keep one copy in a separate place (for example in a different building on the company's site). In the case of hard disc-based systems, two systems should be acquired that are set up in different fire sectors, their data being maintained synchronously.

3.10.5 Logical access control

Various access levels must be set up for access to the archive. These are at least:

- Creator sets up documents in the archive and has reading access to his own document base
- Archivist has reading access to the whole stock of addresses
- Administrator issues user authorizations and has reading access to the document base (for example for backup purposes)

[5] For example EMC Centera, IBM DR-550

[6] For example windream Objectstore

3.10.6 SOPs

The operation of the archive system is naturally subject to the general requirements regarding the operation of computerized systems in the GxP environment. The following regulation sectors should be covered by SOPs:

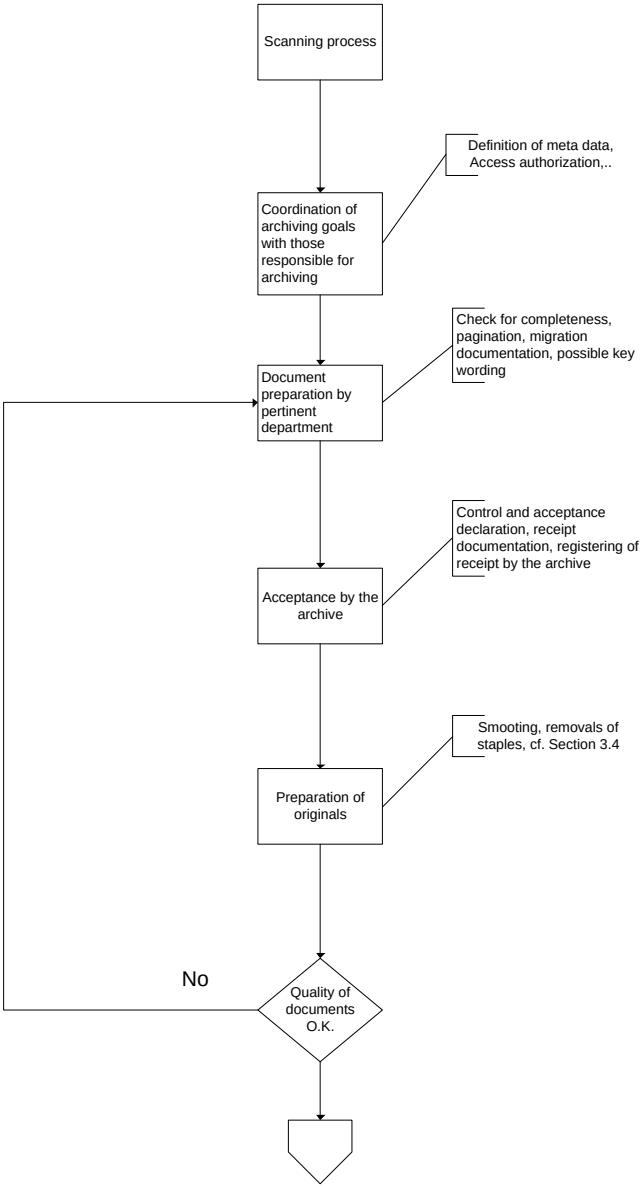
- User administration, issue of user authorizations, cancellation of user authorizations, application procedure and documentation
- Use of the system, importing documents into the archive, and searches
- Administration of the system (for example outsourcing of stocks of data, and regular checking of the synchronization of media, etc.)

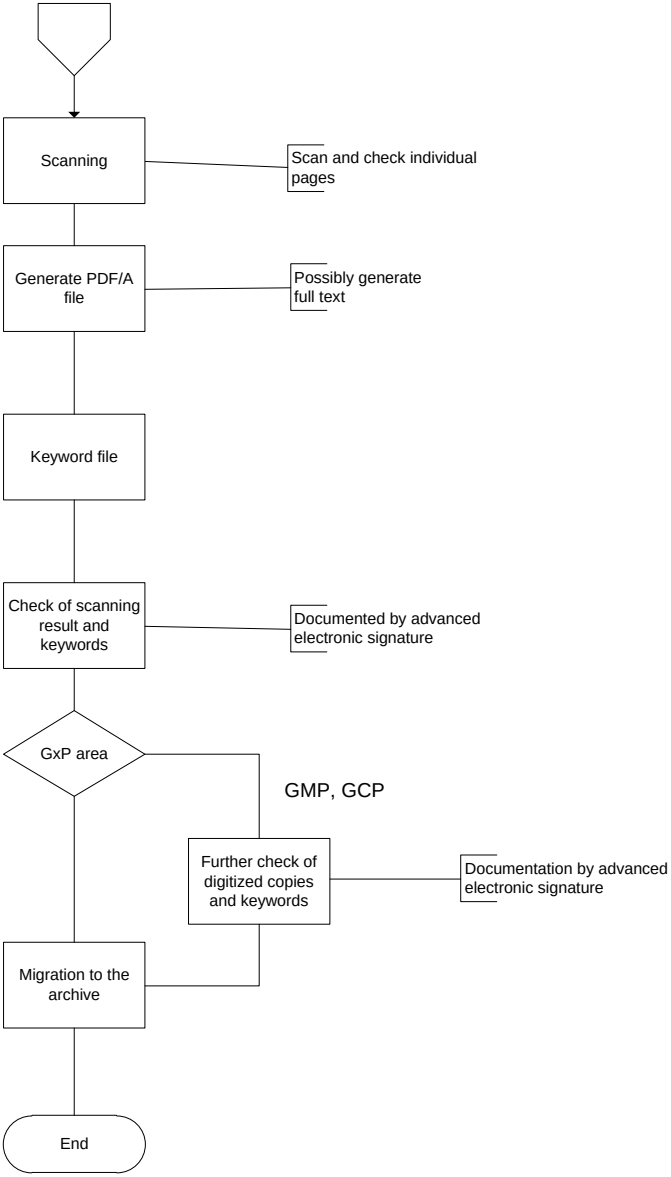
4 References

Heinrich C., Hertlein M., Krull S., Linz T., Opitz U.-A., Schwamberger J. ,Weltmann F.:

Elektronische Archivierung von Papierdokumenten, Pharmazeutische Industrie 2007 69:7 (791-794)

5 Graphic representation of the scanning process




Scanning
Scan and check individual pages
Generate PDF/A file
Possibly generate full text
Keyword file
Check of scanning result and keywords
Documented by advanced electronic signature
GxP area
GMP, GCP
Further check of digitized copies and keywords
Documentation by advanced electronic signature
Migration to the archive
End