

Pharmacogenetic research and data protection – challenges and solutions

J Reischl¹, M Schröder², N Luttenberger³, D Petrov², B Schümann³, R Ternes¹ and S Stürzebecher¹

¹Schering AG, Global Pharmacogenomics, Berlin, Germany; ²TEMBIT Software GmbH, Berlin, Germany and ³Christian Albrecht Universität, Institut für Informatik, Kiel, Germany

The Pharmacogenomics Journal advance online publication, 31 January 2006;
doi:10.1038/sj.tpj.6500372

Introduction

Pharmacogenetic research has been considered a special form of genetics and, therefore, many of the specific ethical, legal and social concerns have been applied to pharmacogenetics accordingly. Consequently, existing guideline documents and legislation on genetics – as well as for biobanking – are also meant to cover pharmacogenetics (e.g. on the international level UNESCO,¹ CIOMS,² European Data Protection Directive,³ CoE Explanatory Memorandum to Recommendation (97) 5,⁴ Working paper of the section 29 Data Protection Working Party⁵ or on the national level 'Biobanks for Research' by the German National Ethics Council⁶). The broad and controversial spectrum of views is reflected in the following citations: The UNESCO paper states: 'Human genetic data have a special status. Due consideration should be given and where appropriate special protection should be afforded to human genetic data and to biological samples'.¹ In clear contrast to this position is the report of the Nuffield Council on Bioethics: 'Given the similarities between genetic and other forms of

personal information, it would be a mistake to assume that genetic information is qualitatively different in some way'.⁷

As pointed out by the Nuffield Council (see above) and in our view, a clear distinction between exploratory pharmacogenetics and genetic diagnostic tests can be drawn, however, once exploratory analyses led to a reliable and validated test for genetic markers for safety or efficacy of drug treatment, their diagnostic value and social and ethical implications may become comparably important to the patient. Therefore, the main question should possibly be if genetic data including pharmacogenetics deserve special consideration at all or whether the focus should be on the informational risk the data pose rather than on their source.

In this ongoing debate, it is our position that both the concerns of the public – including the view of many ethics committee members – have to be acknowledged by an elaborate practice for handling genetic samples and data, and at the same time, an understanding of the benefits and risks of pharmacogenetics has to be promoted. With this in mind, our data protection concept (DPC) for pharmacogenetic studies in clinical research is not meant to indicate that we support 'genetic exceptionalism', but rather that we provide the necessary tools to flexibly react to and be 'compliant' with the public perception and legal

policies, respectively. GENOMatch was designed and implemented in anticipation of even stricter data protection legislation based on the European Data Protection Directive³ and following the examples of, for example, Italy and France. The third important element to be considered obviously is the regulatory view (FDA, EMEA, PMDA/MHWL Japan) in particular with regard to the request for access to original data, including explicitly pharmacogenetic data as stated in EMEA.⁸ In the short run, data protection policies may also become a competitive factor in pharmacogenetic research with regard to investigators' and patients' trust that data will be protected to the best possible extent that is compatible with their optimal research use. Patients may decide about their participation in pharmacogenetic studies based on the understanding of and confidence in data protection measures and the policies how they can potentially benefit from the results of this research.

It is in this context that we developed a detailed DPC for storage of genetic samples and data. The scope of this paper will not include a detailed discussion of ethical and societal questions that are being raised in the context of genetic and pharmacogenetic research, but will focus on data protection highlighting the principles and the most important details of our concept.

Concept

It has been Schering AG's approach for the last 3 years that clinical end points determine study design and sample size and to let clinical results drive the pharmacogenetic questions and related analyses. Pharmacogenetic studies either are integrated into clinical trials as an optional satellite study or are implemented as amendments to already ongoing clinical programs.

The approach of a stepwise implementation of pharmacogenetics into

clinical development is paralleled by a three-step implementation plan for GENOMatch. These steps are (a) 'sample and save' of DNA, (b) 'decision on DNA analysis based on clinical findings and subsequent genetic data storage' and (c) 'biostatistical analysis of clinical and genetic findings = pharmacogenetics'.

To handle, process, and analyze pharmacogenetic samples and data, an IT-based management system called GENOMatch was set up at Schering AG. GENOMatch supports worldwide logistics for pharmacogenetic sampling and biobank administration. GENOMatch is designed to achieve an optimized balance between confidentiality of genetic information and the informational risk for volunteers participating in pharmacogenetic trials on the one hand and the requirements of the researcher to accurately link clinical and genetic data on the other. It was developed in close collaboration between Schering AG, Tembit Software GmbH, and the University of Kiel, Germany.

The GENOMatch design aims at separation of information and responsibilities via a restricted access right approach based on the respective role and institution and was determined by the following key requirements:

- GENOMatch must support the smooth integration of pharmacogenetic research processes into already existing clinical development processes.
- Feedback of relevant information to the patient and withdrawal of sample and data must be possible at any time during sample storage (requiring pseudonymization, not allowing anonymization).
- GENOMatch must be prepared for future data protection regulation and legislation.

Based on the principles of (a) pseudonymization by triple-coding (see Table 2 for list of technical terms), (b) role specific access to data (controlled by a cryptographic chip card system), (c) encryption of all data being transferred, a systematic approach of separation of responsibilities has

been implemented in the following way:

- separation of responsibilities *by institution* with different institutions being in charge of
 - keycode list storage (see list of technical terms): 'Secure Identity Management' center = SIM center; hosted by an independent and reliable application service provider (Dataport)
 - sample storage: 'Central Sample Repository'
 - genetic (sequencing) analyses: 'Analysis Lab'
 - combination of clinical and genetic data (pharmacogenetic analyses) in a 'Secure Data Area'
- separation of responsibilities *within institutions* (central lab and sponsor) regarding
 - independent steps of triple coding with change in personnel responsible for the coding and storage procedure at the 'Central Sample Repository'
 - no access to the original ('identifiable') clinical data set for the personnel performing pharmacogenetic analyses in the 'Secure Data Area'
 - handling of data sets necessary for pharmacogenetic analyses by designated personnel only ('Secure Data Area')
 - only aggregated (see list of technical terms) data are released from the 'Secure Data Area' eliminating the possibility to trace back an individual patient

The DPC was audited internally and externally. It was approved by the Pharmacogenetics Advisory Board overseeing Schering's policies and SOPs in pharmacogenetic research. The DPC was also submitted to the independent data protection authority (Unabhängiges Landesamt für Datenschutz) of Schleswig-Holstein, Germany, and received a certificate of compliance with current data protection law. Furthermore, the data protection authority judged the GENOMatch concept qualified to satisfy future legislation.⁹

All project partners who either are occupied with concept implementation or take over roles in the GENOMatch process (see Figure 1) have been ISO 9000 (or comparable QM system) certified for several years. The project was managed within the Schering quality management system and was successfully validated regarding Good Clinical Practice standards,¹⁰ Code of Federal Regulations, Chapter 21, Part 11¹¹ and the Data Protection Concept.¹²

GENOMatch implementation

GENOMatch represents a flexible (with regard to customization) internet-based application that makes future extensions, maintenance, and support competitive. To combine the different demands of the involved institutions, we decided to follow the 'rapid prototyping' (see Table 2 for list of technical terms) development method. Users were intensively involved in the application development. They had access to the system to get used to the workflow and to assure a work environment compatible design of GENOMatch.

The GENOMatch IT system is comprised of several modules with a three level architecture as shown in Figure 2.

The first level is represented by a browser-based (see Table 2 for list of technical terms) user front end. To ensure the role-based concept, each user gets access to the system with a smart card and personal identification number (PIN). To control user access to the GENOMatch IT system, a role-based access dispatcher (see Table 2 for list of technical terms) system is applied.

The Secure Identity Management (SIM) center represents the second and third level of the GENOMatch IT system. It is hosted by the independent institution Dataport mentioned above (application service provider). The second level holds the indicated applications: access dispatcher, identity management, and sample tracker application. The third level holds the related databases.

The complete organization of handling pharmacogenetic samples and

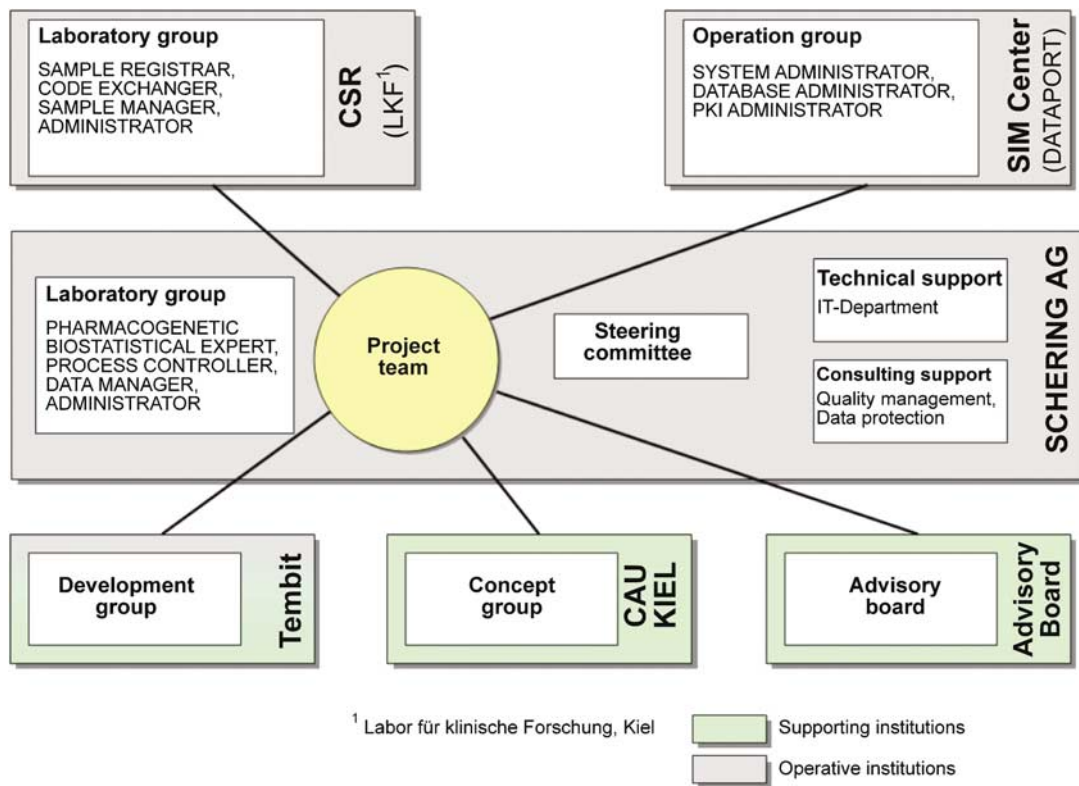


Figure 1 GENOMatch Working Groups. Description of the GENOMatch project structure listing all GENOMatch partners and roles. Each role was represented from the very beginning in the project to ensure user friendliness and practicability of the GENOMatch workflow.

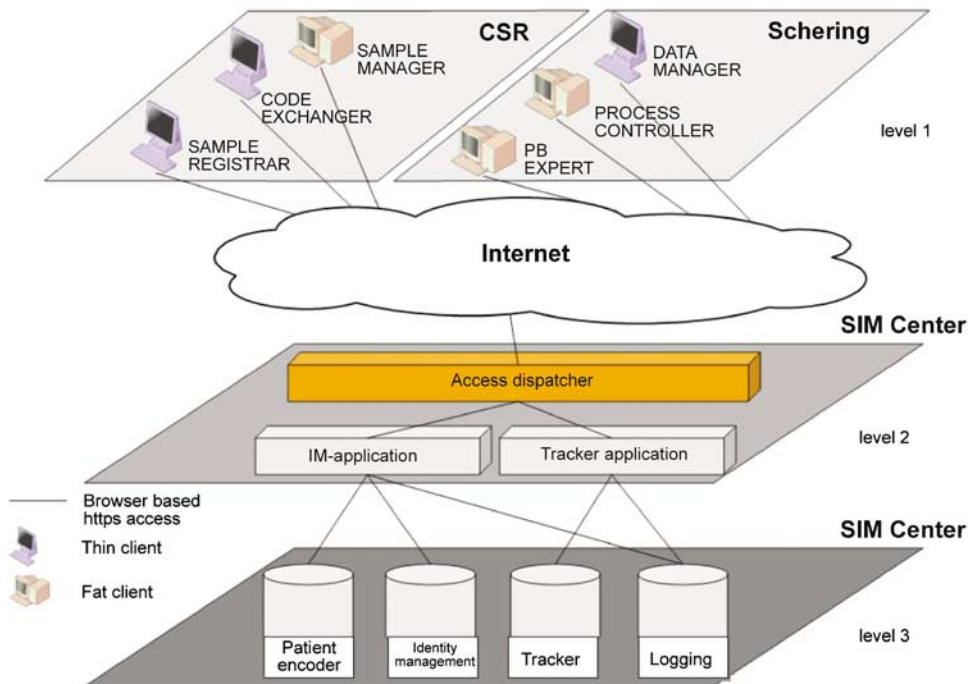


Figure 2 GENOMatch system architecture. The GENOMatch system consists of three levels. Level 1: user front end reflecting the role based concept of GENOMatch. All users communicate via the internet (https). Level 2: GENOMatch application. The access dispatcher ensures that only authorized users get access to their role specific function. Level 3: GENOMatch databases. The data-model ensures a strict separation of keycode list administration (patient encoder and identity management), tracking information (tracker) and audit trail (logging).

data are represented and safeguarded by the GENOMatch process either by the application itself or via Standard Operation Procedures (SOPs). A brief overview of the interplay of SOPs, the DPC and the GENOMatch application is shown in Table 1.

The following sections describe the major functions/institutions of GENOMatch necessary for 'tracking' a sample and its associated data from sample collection in the clinic (study center) to the aggregated result of pharmacogenetic analyses.

Central sample repository

Once a sample for pharmacogenetic analysis has been collected, it is labeled with two different identifiers, namely the Patient Number PN (a unique patient pseudonym in a given

study) and a barcode 1 (BC1). The sample is sent to the *Central Sample Repository* (CSR) together with a patient accompanying letter (PAL) containing additional plausibility control information (e.g. patient's date of birth) and a confirmation by the investigator that Informed Consent has been obtained for participation in the pharmacogenetic study.

The CSR checks the plausibility control information and the informed consent confirmation, removes the Patient Number and barcode 1 (BC1) from sample tubes, and stores the tubes using a barcode 2 (BC2) that can no longer be associated with the PN within the CSR. GENOMatch users operating at the CSR are: the *Sample Registrar* (SR), the *Code Exchanger* (CE),

and the *Sample Manager* (SM) (see Figure 3).

SR and CE must be different persons as determined by a related SOP. They perform two pseudonymization steps (triple coding; see Table 2 for list of technical terms) such that the sample that is handed over to the SM – as a result of triple coding – can be considered factually anonymized (see list of technical terms): The SR removes the patient number, archives the PAL, and hands over the sample to the CE; the CE replaces the first barcode by a new barcode (BC2). Due to the fact that the system must allow several samples to be drawn from the same patient, the BC2 labels of these samples are grouped by the SIM Center (via the patient number) under a *Sample Group Number* (SGN). To allow the generation of a keycode list for any given study, two kinds of sample identifier pairs per sample are transmitted to the *Secure Identity Management* (SIM) center by different persons in the CSR: (1) PN–BC1, and (2) BC1–BC2. The SR and CE operate through Thin Client computers. These devices provide only functions needed to conduct the steps of sample handling assigned to the respective roles, and are configured such that no other programs can be installed. The combination of smart card login and Thin Client configuration restricts the possibility of fraudulent use. This is especially important for the SR role, because this user deals with patient identifying data (both for assignment and withdrawal).

An additional CSR task is to track samples during their entire life-time of storage and analyses. The CSR provides information on, for example, number of samples, samples sent for DNA analysis, amount of DNA available, etc. Sample tracking information is also stored at the SIM Center.

The SM is in charge of storage and retrieval of the samples in case of analysis or withdrawal and is acting according to requests from the *Pharmacogenetic Biostatistical Expert* (PBE).

We have so far established a single CSR, but GENOMatch allows further CSRs to be added and to function as a

Table 1 Interplay of data protection concept, SOPs, and GENOMatch IT system (main component(s))

	DPC	SOP	IT
<i>General tasks for all institutions</i>			
User qualification to handle the system		✓	
Workplace access and equipment	✓	✓	
System access	✓		✓
User management	✓		✓
<i>Schering only</i>			
Study management	✓		✓
Integration of supporting LAN applications		✓	
Access to supporting LAN applications			✓
Sample management	✓		✓
Study control and supervision	✓		✓
Handling and preparation of clinical study and genomic data	✓	✓	
Combination of clinical study and genetic data			✓
<i>Central Sample Repository only</i>			
Pseudonym (Barcode 1 and 2) management	✓		✓
Patient assignment	✓	✓	
Patient management	✓		✓
Workflow for pseudonymization (triple-coding)	✓	✓	
Sample management	✓		✓
<i>Dataport only</i>			
System maintenance		✓	
Public key infrastructure (PKI) management	✓		✓
Smart card shipment		✓	
System update			✓
System operation surveillance			✓
<i>Tembit only</i>			
Application release and supply		✓	
Application upgrade			✓

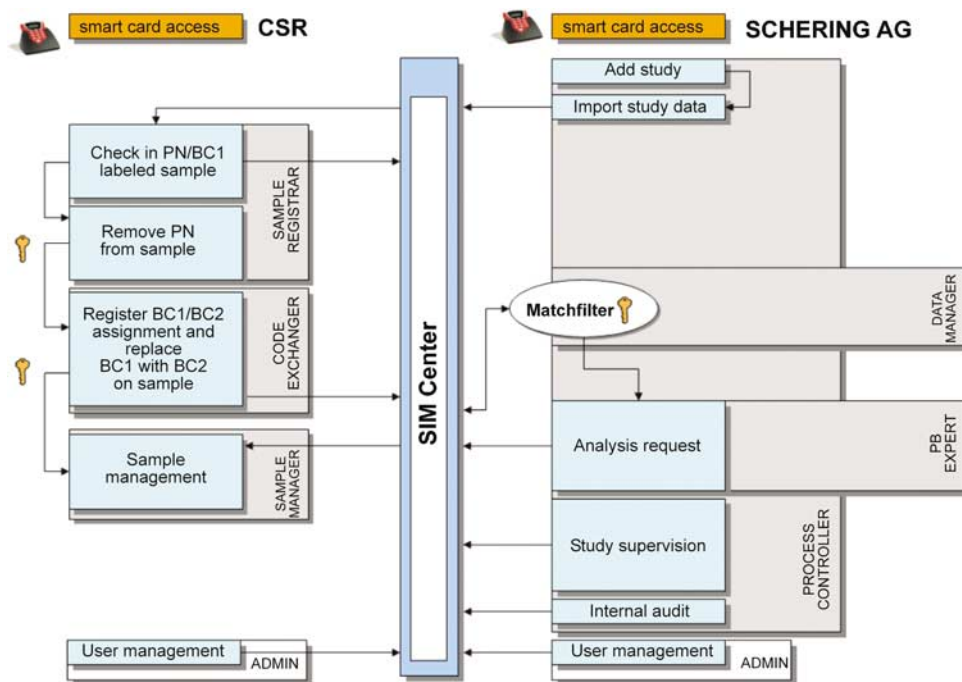


Figure 3 Characteristics of the GENOMatch workflow. The SIM-Center acts as a hub, where all information is collected and provided to the users according to his/her access rights and roles. Section GENOMatch implementation describes the workflow in detail.

single virtual sample repository. The CSR(s) together with the SIM Center (see below) represent the DNA bio-bank.

Secure data area

The infrastructure for pharmacogenetic analysis based on individual genetic data and clinical data is located in the *Secure Data Area* (SDA). These analyses are performed under a high degree of additional security measures to exclude any leakage or misuse of data. All data being processed in the SDA are identified by a BC2 identifier. All samples of the same patient can be accessed through the SGN as described above.

The roles in the SDA are: the *Process Controller* (PC), the *Pharmacogenetic Biostatistical Expert* (PBE) and the *Data Manager* (DM). They are operating in an access-controlled room.

The PC can initiate new GENOMatch users, new CSRs and new studies. He or she is in charge of supervising the process by analyzing audit trail information generated automatically by the GENOMatch system. This access is restricted in a way that the data protection principles cannot

be circumvented, that is, access is only granted to workflows, but not to individual data entries.

The genetic analysis strategy is driven by the outcome of the clinical study. The PBE selects (supported by the sample tracking system of the CSR) the samples of interest together with the appropriate extraction/analysis technologies. This information is transmitted directly to the Sample Manager of the respective CSR, who processes this request (i.e. sending the samples to the extraction/analysis labs). Genetic data being generated in analysis labs will be sent directly to the SDA. Clinical data records of selected patients are transferred into the SDA after passing through a so-called Match Filter (see Section The Match Filter).

The DM takes care of loading genetic data into the genetic database. By preparing the genetic data set and the clinical data set, the DM enables the PBE to carry out the statistical analysis, that is, matching genetic data with clinical data. The aggregated result of such an analysis will then be reported and stored outside the SDA. Individual genetic data will remain stored in the

SDA, clinical data will be stored only temporarily during the time of pharmacogenetic analyses. Aggregation of data will be performed at a level where identification of individual results can be ruled out.

Secure identity management center

The SIM Center is the core of the GENOMatch IT system. It is the only place where the keycode list for PN, barcodes (BC1 and BC2) and SGN is stored. Each function mentioned above can only retrieve restricted information from this keycode list. The PBE, for instance, can only retrieve BC2s for patients with available genetic samples.

Since the keycode list is essentially the core of the system that allows the linkage of clinical data with genetic data, an elaborate backup strategy to prevent data loss while retaining pseudonymization is essential and is provided by the SIM Center.

For audit-trail purposes, all actions carried out by GENOMatch users are logged and stored in the logging database that can be accessed by the PC. In order to administrate the bio-bank, the SIM Center also stores the

Table 2 List of technical terms

Aggregated data: Summary data after statistical analysis that do not allow to reconstitute individual contributions leading to the overall result (exception: in case of 100% association of genetic information to a certain phenotype)

Browser: A web browser is a software application that enables the user to view and exchange data via the internet. Since GENOMatch is an internet-based system, the user front end is 'browser based'

Anonymization: Anonymization of data means destroying the link between sample and donor (irreversibly de-linking data)

Factually anonymized: In the case of GENOMatch, the keycode list allows to link genetic information to clinical information of individual donors. However, due to the multi-step coding procedure, it would require an excessive amount of time, expenses and manpower to circumvent the data-protection measurements implemented in GENOMatch. Hence, the term 'factually anonymised' is used

Keycode list: List being stored in encrypted form at the SIM-Center where identifier (such as PN, BC1, BC2, etc.) relationships are kept. In GENOMatch the 'patient encoder' table constitutes the keycode list, the identity manager administrates the barcode history

http: HyperText Transfer Protocol (http) is a standardized protocol to transfer information via the internet. https is a secure version of http using cryptographic protocols to secure communications on the internet

Pseudonymization : The pseudonymization (coding) replaces the patient identifier (e.g. patient number) with a new identifier (barcodes in the case of GENOMatch). Link to the respective individual is possible but aggravated. In GENOMatch a triple coding procedure is implemented

Rapid prototyping: Software engineering methodology used to develop GENOMatch. Rapid prototyping that relies on several prototyping/refinement cycles as opposed to strictly sequential methods, where all user requirements have to be defined in advance

Smart card: Plastic card (in the size and shape of a credit card) with embedded cryptographic processor and a personal certificate. The smart card is issued to a single person together with a numeric password (PIN). Only the combination of possessing the smart card and knowledge of the PIN allows access to the role specific functions at the GENOMatch SIM-Center

Role-based access dispatcher: Software module that implements the GENOMatch role based concept by ensuring that every user logging into GENOMatch (authorized by smart card and PIN) gets ONLY access to the function assigned to his/her role

Triple coding: Coding procedure for genetic samples that comprises three steps: The first code, the patient number (PN) is replaced by the Barcode 1 (BC1), which again is replaced by the barcode 2 (BC2). This multi-step procedure conducted by different persons ensures that no single person in the GENOMatch process is able to reconstruct the complete key code list or any single patient identity

Web service: Software system that supports machine-to-machine interactions via the internet

sample status information, which is updated by the SM and can be accessed by the PBE.

The SIM Center also provides aggregated information, which allows assessing the current recruitment status of

patients consenting to pharmacogenetic analysis.

Technical features. Users have access to the GENOMatch internet application via https (see Table 2 for list of

technical terms). In spite of the internet use, the system is seen as a 'closed' application because all users must be known to the system.

The GENOMatch *Access Dispatcher* (see Table 2 for list of technical terms) (cf Figure 2) checks all incoming access requests. As mentioned above, it is built on an elaborate role-based access control concept. To enforce this concept, users need a valid smart card (see Table 2 for list of technical terms) and a personal identification number (PIN) for authentication. After successful authentication, the *Access Dispatcher* controls the user authorization to access the *Identity Management* (IM) application and/or the *Tracker* application. Four databases represent the third level of the GENOMatch IT system. Patient-related information is stored in two separate databases called 'patient encoder' and 'identity management'. The patient encoder database holds the assignment of patient number, study number, and SGN. The identity management database is responsible for the barcode administration including barcode generation, 'history', and assignment. Both, patient encoder and identity management are accessed only by the CSR roles. The data in both databases are encrypted to secure them even in case of a hardware theft.

The 'tracking database' stores information on sample tracking, recruitment and user related data. User data are processed by the Access Dispatcher to decide what kind of access the current user has.

Finally, the logging database logs all GENOMatch actions to enable auditing of the GENOMatch process by the PC or by external auditors.

GENOMatch uses external sources to upload clinical data and study related information (e.g. trial site information). The integration of external sources is approached by converting the different data formats (e.g. SAS-files) into XML, a standardized data exchange format. Furthermore, web services (see list of technical terms) are used for communication between system components and external application like the local sample tracking system of the CSR.

The Match Filter. The Match Filter application is designed to upload clinical data into the SDA in order to reconcile clinical and genetic data without compromising data protection. It has three tasks:

- dropping the clinical data records from patients that did not participate in the pharmacogenetic sub-study;
- cleaning clinical data records from the personal data (e.g. initials)
- coding clinical data with the same barcode 2 as the corresponding genetic data.

The pseudonymization procedure is carried out in two steps by the Match Filter application. This is reflected by its modular architecture.

The first module is operated by the DM outside the SDA. The DM receives input data from clinical result sets. During this step, the PNs are exchanged to Match IDs, which are generated and stored at the SIM-Center. The Match Filter transfers only PNs and the study number via the Internet. Furthermore, initials are dropped and the date of birth is replaced by age.

The second module is operated by the PBE inside the SDA. In this step, Match IDs are replaced by the corresponding BC2s and all patient data sets with no corresponding genetic sample will be dropped. Multiple tubes for each patient will be represented in a new 'matchtable'-file. Thereby each tube can be linked to a complete clinical data set. Throughout the whole process, no clinical data leave the SAG environment.

On basis of BC2 identifiers, clinical and genetic data can be matched and statistically analysed.

Testing

The system has been tested thoroughly. A series of white label studies were executed prior to formal validation to test the usability and performance of the prototype system. After the prototype met the expectations, a formal validation was executed. Validation testing was based on documented user requirements, which had been risk-categorized regarding compliance

to Good Clinical Practice standards,¹⁰ Code of Federal Regulations, Chapter 21, Part 11¹¹ and the Data Protection Concept.¹² The tests included disaster and stress tests (e.g. disconnect system from power, huge amount of data input, and using inappropriate data types as input) in addition to functional tests.

During the validation process, a reference system (identical to the production system) was set up and tested on the base of automated, script-based regression tests with more than 1 million barcodes (BC1 and BC2).

A framework for maintenance, operation, and change management of the system has been set up and formalized in respective directives (e.g. SOPs). External partners were audited as required, to ensure adequate qualification.

Feedback of individual genetic data to patients

The question of individual return of genetic data from pharmacogenetic studies has been a topic of controversial discussion, in particular with regard to the aspect that most information will be non-validated and that genetic counselling cannot be provided for in large clinical studies. Beyond the proponents view on return of genetic data (e.g. in ^{1,2,4}), we considered it necessary to be prepared for situations in which return of genetic data may be of substantial relevance to the patient and the treating physician, for example, if a certain treatment is found to be of high risk and low or no benefit. In such cases, protecting the patient from harm – even if the harm will only occur with a certain probability – should take precedence over the justified demand for validation of the assay and other concerns regarding the premature disclosure of genetic data. There are, of course, other instances where the sample donor may explicitly request feedback of genetic information (see below). It is for these reasons that we have designed a feedback process that allows the return of individual genetic

information without interfering with the data protection policy of GENO-Match.

The approach described below is one possibility to implement individual feedback of genetic data and is focused on maintaining the high level pseudonymization applied by the GENO-Match sample and analysis functionality. Less strict concepts to implement feedback might be appropriate.

Before any such feedback of genetic information is given, a complete re-run of genetic analyses according to the procedure described for the primary sample will be performed, that is, only confirmed results will be forwarded to the physician/the patient.

In order to maintain the GENO-Match data protection principles also in case of individual feedback, a separate institution – the Feedback Handling Center (FHC) – will be implemented (see Figure 4). Its main task is to release password protected individual patient data. A 'Feedback Number' will be generated by the SIM center early in the process. The Feedback Number is unique and will provide the necessary information for each of the roles in the feedback process selectively, for example, the address of the physician of choice for the SR and the FHC or the SGN for the PBE.

In a first step, the FHC will provide the investigators with aggregated results of a pharmacogenetic study together with a *Password 1* (PWD1) per patient. The investigator informs the patient about these results. Based on these findings, the patient can decide whether to request feedback. In case of a feedback-request, the patient can choose either the investigator or any other physician. The physician chosen by the patient authorizes himself with the PWD1 and a second informed consent where the patient confirms his or her interest in individual feedback at the FHC. A successful authorization triggers the sample registrar to equip the authorizing physician with a blood collection tube labeled with a BC3, which indicates a 'validation sample'. Furthermore, the feedback requesting patient is equipped with a PWD2 via the physician of choice. The physician of

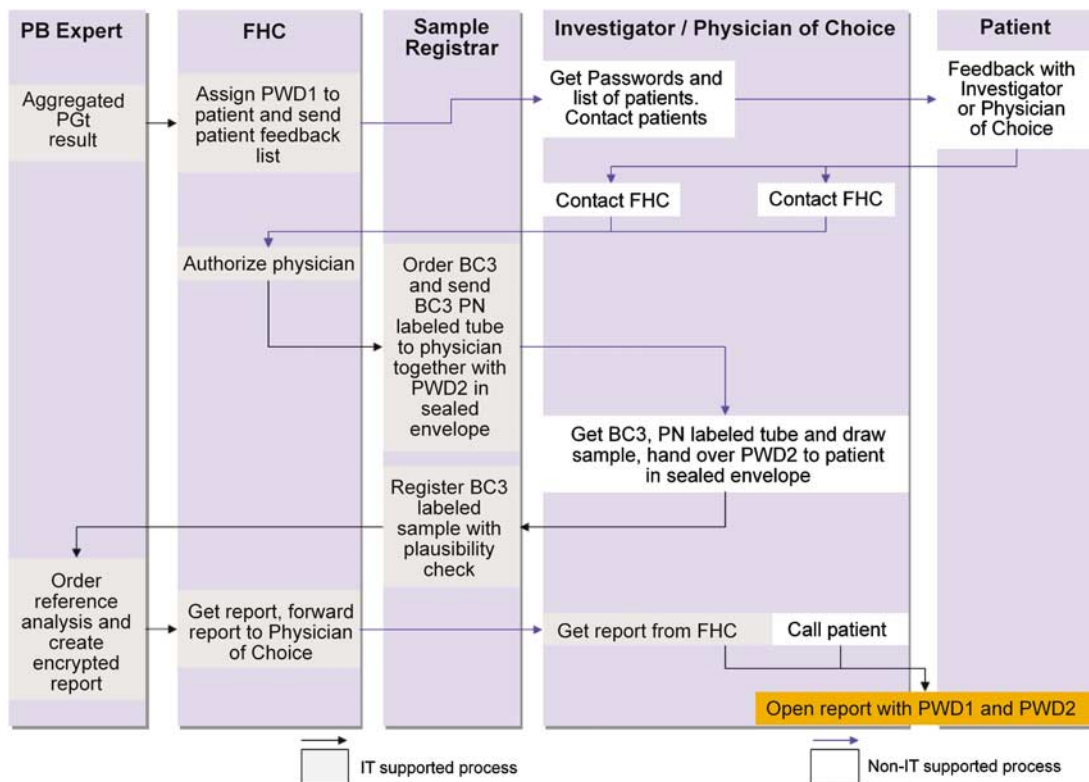


Figure 4 The GENOMatch feedback process depicting the crucial steps. The FHC acts as a firewall between PBE and the investigator/physician of choice to ensure confidentiality. The patient is only contacted by the investigator/physician of choice. In order to access individual genetic data both the patient and the investigator/physician of choice have to cooperate. For details see Section Feedback of individual genetic data to patients.

choice draws the validation sample and sends it to the SR at the CSR, where it is processed as described above in the GENOMatch system. Checking in a BC3 labeled sample triggers the SIM Center to inform the PBE that a re-analysis with this validation sample has to be carried out in order to confirm the findings for a particular SGN data set. Only after replication of findings, an individual feedback report will be generated by the PBE and encrypted in a way that both passwords, the password of the physician (PWD1) and the password of the patient (PWD2), are necessary to access the information. This ensures that only physician and feedback requesting patients together can read the document. The encrypted data set is transferred via the FHC to the respective physician. The patient will obtain the requested information without his or her identity being revealed to any institution represented in the GENOMatch process.

The implementation of the feedback procedure has been completed and

tested. The concept is currently under review by the data protection authority ULD. The next step will be the transfer of the IT tools and SOPs into the production system.

Conclusion

Although we consider many of the special measures of pharmacogenetic sample and data handling as a ‘necessary compromise’ that will most likely be overcome by the future development of pharmacogenetic research and the better understanding and perception of its benefits and risks in the public, a number of principles will most likely be maintained that may – in the long run – have repercussions on the handling of clinical research data as well.

The debate about the justification of ‘genetic exceptionalism’ will likely lead to reconsideration of the level of data protection for all sensitive medical data. This will also help to differ-

entiate medical data by their potential ‘informational risk’ rather than generally stigmatizing genetic data as particularly sensitive.

Among the special issues of pharmacogenetics, the option of feedback of relevant and validated information to the patient is debated controversially with regard to its compatibility with data protection measures, the potential uncertainty it may cause for the patient and the appropriate provision of counseling.¹³ The availability of aggregated data to doctors and patients appears to be of less controversy. We have designed the GENOMatch system to allow feedback to patients without jeopardizing the established DPC and are able to accommodate the requests of patients and doctors as well as the development of guidelines and legislation. The acceptance of the GENOMatch concept will depend on the sponsors’ and investigators’ ability to explain and make transparent the implications of this system to Ethics Committees and patients.

The orientation of GENOMatch towards current and future legislation – as far as it can be ‘predicted’ – provides an optimal basis for adjustments and modifications while maintaining the main functionalities as designed and implemented today.

Acknowledgments

We thank Dr B Müller (Schering AG) and Professor Lutterbeck (TU Berlin) for his valuable advice and continuous support of the project.

Correspondence should be sent to Dr S Stürzebecher, Global Pharmacogenomics, Schering AG, Müllerstrasse 178, 13342 Berlin, Germany.
Tel: + 49 30 46811354
Fax: + 49 30 46811323
E-mail: ClausSteffen.Stuerzebecher@SCHERING.DE

References

- 1 UNESCO International Bioethics Committee. DRAFT International Declaration on human genetic data, Addendum 2, 8.10. 2003.
- 2 Council of International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO). International Ethical Guidelines for Biomedical Research Involving Human Subjects, 2002.
- 3 European Parliament and Council. European Parliament and Council: Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal L281, 23/11/1995, 31–50.
- 4 Explanatory Memorandum to Recommendation No. R (97) of the Committee of Ministers to Member States, Adopted by the Committee of Ministers on February 13, 1997. Available at: [http://cm.coe.int/ta/rec/1997/ExpRec\(97\)18.htm](http://cm.coe.int/ta/rec/1997/ExpRec(97)18.htm).
- 5 Article 29 data protection working party, 1 August 2003 12168//02/EN WP 80 available under: http://www.statewatch.org/news/2004/feb/biometric-wp80_en.pdf.
- 6 German National Ethics Council, Biobanks for Research, 2004; Available under <http://www.ethikrat.org>.
- 7 Nuffield Council on Bioethics. Pharmacogenetics ethical issues 2003.
- 8 EMEA. Position Paper on Terminology in Pharmacogenetics, 21 November, 2002 (EMEA/CPMP/3070/01).
- 9 Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein: brief report on the data protection audit data processing infrastructure concept of the Schering corporation for the secure pseudonym storage and keeping of blood and tissue samples intended for genetic analyses. Available under: http://www.datenschutz-zentrum.de/audit/kurzgutachten/a0303/a0303_engl.htm.
- 10 ICH Topic E6. Guideline for Good Clinical Practice, CPMP/ICH/135/95 1997.
- 11 FDA. Department of Health and Human Services, 21 CFR Part 11 1997.
- 12 Luttenberger N Data protection concept for the ‘Sample and Save’ Part of the GENOMatch Project at Schering AG, Ver. 2.0 (available upon request).
- 13 Renegar G, Webster CJ, Stürzebecher S, Harty L, Ide SE, Balkite B *et al*. Returning genetic research results to individuals: points-to-consider. *Bioethics* 2006; **20**: 24–36.