Fondazione Smith Kline

Guidelines
for Clinical Protocols in Genetic Research

Recommendations for the drafting and assessment of clinical research protocols in genetics

AUTHORS

(* An asterisk after an author's affiliation indicates that s/he represented the organization concerned within the working group which drafted the present Guidelines.)

Teresa Annecca

National Data Protection Office, Rome.

Adriano Bompiani

President, Italian Society for Bioethics and Ethics committees* ;
Council of Europe Steering Committee for Bioethics (CDBI).

Roberto Buttura

President, the Citizens' Movement*.

Giliola Calori

Statistician, Clinical Research Office, H. S. Raffaele del Monte Tabor Research Institute ;
Ethics Committee, H. S. Raffaele del Monte Tabor Research Institute.

Marisa Dell'Aera

Vice-President, Ethics Committee, City of Bari Hospital Board*.

Annarita Meneguz

Director, Biochemistry, Pharmacology and Coordination of Technical and Scientific Consultancy, Department of Pharmacy, *Istituto Superiore di Sanità* ;
Italian expert at the EMEA Safety Working Party .

Demetrio Neri

Professor of Bioethics, University of Messina ;
Co-editor, *"Bioethics – an Interdisciplinary Journal"*.

Giuseppe Novelli

Professor of Medical Genetics, University of Tor Vergata, Rome ;
Italian Society of Human Genetics* ;
EMEA Committee of Pharmacogenetics (London).

Pier Franco Pignatti (Coordinator of the working group)

Professor of Molecular Genetics, University of Verona ;
Past President (2002-2005), Italian Society of Human Genetics* ;
President Elect, European Society of Human Genetics.

Antonella Pirazzoli (Secretary of working group, responsible for drafting the present document)

Head of Clinical Genetics, GlaxoSmithKline S.p.A. ;
Coordinator, Genetics Working Group, Smith Kline Foundation*.

Giuseppe Recchia

Medical Director, GlaxoSmithKline S.p.A.*.

Paolo Rizzini

Vice-President and Secretary General, Smith Kline Foundation*.

Antonio G. Spagnolo

Professor of Bioethics, University of Macerata ;
Ethics committee, A. Gemelli University Hospital, Rome*.

Pasquale Spinelli

President, Federation of Italian Medical-Scientific Societies*.

Giampaolo Velo

Professor of Pharmacology, University of Verona ;
Director, Medical Pharmacology Unit ;
Director, Reference Centre for Education and Communication within the WHO program for International Drug Monitoring ;
Coordinator, Section of Clinical Pharmacology, Italian Society of Pharmacology.

CONTENTS

TABLES

FIGURES

TEXT-BOXES

PRESENTATION

We are pleased to present the volume *Guidelines for Clinical Protocols in Genetic Research - Recommendations for the Creation and Assessment of Clinical Research Protocols in Genetics.*
This manual is intended as a possible help for researchers who wish to prepare a human genetic research project and for the ethics committees such projects will be submitted to.
Thanks to the Human Genome Project and the availability of powerful new methods of genetic analysis, more and more applications are submitted for genetic research, or for clinical research involving collection of biological samples for human genetic analysis and handling of genetic data. This trend is illustrated in the tables at the beginning of the present volume.
In our opinion, there was thus a need for a text such as this, based on discussions among a team of experts; it comes at the right time to provide information and up-to-date knowledge that were not readily available before. The multidisciplinary group which drafted the present guidelines was representative of those most directly involved in this field in Italy – bringing together areas of expertise such as human genetics, bioethics, medicine and pharmacology alongside representatives of citizens, ethics committees, institutions, scientific societies and the pharmaceutical industry. The working group's names are listed at the very beginning of this booklet. The Italian Society of Human Genetics and the Smith Kline Foundation contributed to the conception and organization of the project.
The work of the group lasted about two years, involving plenary meetings and e-mail correspondence. The most important scientific and regulatory updates, at Italian, European and international level, have been included. Contents range from general topics, such as the nature of genetic information and genetic tests, to more specific matters, such as the evaluation of risks and benefits and data protection. There are also chapters on study rationale, aims and design, sample size, sample identification, use and handling of biological samples, access to research results, commercial and patent rights, insurance cover, research in minors and mentally disordered subjects, and informed consent.
The publication's aim of providing clear, practical information is reflected in the "Summing up and recommendations" section at the end of each chapter. For ease of reference, these are all included in the opening chapter "Summary and Recommendations". The text also contains two check-lists for the creation and assessment of the research protocol and informed consent, text-boxes highlighting specific topics, tables and figures, a brief glossary of genetic terms used in research applications, and a selected bibliography.
This volume will be distributed to all those potentially involved in human genetic research and can be obtained free of charge from the working group's coordinator. Suggestions and constructive comments submitted to the websites of the Italian Society of Human Genetics (http://sigu.accmed.org) and Smith Kline Foundation (http://www.fsk.it) are welcome and will be taken into consideration for a possible revised edition in the near future.

With our best wishes!

Verona and Milan, October 2006

Pier Franco Pignatti
Italian Society of Human Genetics

Paolo Rizzini
Smith Kline Foundation

The guidelines: aims, methods and scope

Human genetic research in Italy

The advances in knowledge of the human genome during the last decade have paved the way for extensive application of genetic and genomic techniques in diagnosis, prevention and treatment of the most important diseases. Against this background, there have been a number of human genetic and pharmacogenetic studies, starting from the second half of the nineties.
At the time of writing this field of research accounts for 15-20% of protocols submitted to Italian ethics committees, and their number is likely to continue growing rapidly. This trend is illustrated in recent data kindly supplied by the City Hospital Board Ethics Committees of Verona and Bari (see Tables 1 and 2).

Table 1: *Genetic research protocols submitted to th e City of Verona Hospital Board Ethics Committee, 2003 -2005.*

Year	Total protocols	Genetic Protocols*	Pharmacogenetic Protocols*	Total genetic and pharmacogenetic protocols*
2003	64	8 (12%)	3 (5%)	11 (17%)
2004	146	15 (10%)	9 (6%)	24 (16%)
2005	115	13 (11%)	11 (10%)	24 (21%)

** Protocols with both genetic and pharmacogenetic aims have been included in both categories.*

Table 2: *Genetic research protocols submitted to the City of Bari Hospital Board Ethics Committee, 2004 -2005.*

Year	Total protocols	Genetic Protocols*	Pharmacogenetic Protocols*	Total of genetic and pharmacogenetic protocols*
2004	146	5 (3%)	6 (4%)	11 (7%)
2005	109	18 (17%)	12 (11%)	30 (28%)

** Protocols with both genetic and pharmacogenetic aims have been included in bot h categories.*

Despite the high percentage of genetic protocols submitted to ethics committees, knowledge of genetics is still relatively limited and there is a lack of practical guidance for those intending to work in this new field. In particular, issues such as information on DNA samples, data protection and research-related risks are often the focus of queries and required amendments in ethics committee reports.
These trends are highlighted in data kindly supplied by GlaxoSmithKline, showing reasons for which ethics committees rejected, queried or required amendments to the company's genetic research protocol submissions.

Table 3: *Queries and required amendments to genetic research protocols submitted by GlaxoSmithKline, 2004 -2005.*

Year	Total applications	Number (%) of approvals	Number (%) of rejections	Number (%) of queries & required amendments	Number (%) of approvals & rejections after explanations & amendments.
2004	145	107 (95%)	6 (5%)	25 (22%)	Approvals: 22 (88%) Rejections: 3 (12%)
2005	63	54 (95%)	3 (5%)	16 (28%)	Approvals: 9 (83%) Rejections: 2 (17%)

Table 4: *Reasons for rejections, queries and required changes in response to genetic research protocol applications by GlaxoSmithKline, 2004 -2005.*

Year	Reason for rejection (number of occurrences in brackets)	Issues giving rise to queries and required amendments (number of occurrences in brackets)
2004	Investigator not available for genetic research (3) Research not scientifically valid (1) Polymorphisms for analysis not described (1) Genetic results in private sponsor database (1)	Risks should be better detailed (6) More information needed about DNA samples (6) More information needed about background and aim of the study (5) More information needed about data protection (3) Candidate genes should be listed (3) Specify genetic research among the aims of the main clinical study (3)
2005	Data protection (1) Samples stored beyond the end of the study (1) Reason not provided (1)	More information needed about DNA samples (11) More information needed about data protection (8) Candidate genes should be listed (1) More information needed about background and aim of the study (1)

The need for shared standards

In Italy, the need to establish standards for genetic research and indicate the fundamental principles that should inspire such research was first stated in the "Declaration of Erice on the ethical principles of pharmacogenetic research" [(1)], issued in March 2001; the most exhaustive and specific document to date on pharmacogenetic research is the "Proposed guidelines for assessment of pharmacogenetic experimentation" [(2)] (November 2001), drafted by a multidisciplinary team of experts in a joint action by GlaxoSmithKline and the Italian Society of Hospital Pharmacy. Standards for correct management of genetic research in humans are far from being defined and shared. Against this background, the present guidelines aim to help define such standards for scientists contributing to clinical study design in a genetic research setting, for members of the ethics committees to which the resulting research proposals are submitted, and for all those involved in genetic research applied to the improvement of human health.
Other recent or forthcoming documents may further explain various matters related to genetic research, and help the development of specific standards of good practice within this field. The following list includes any such documents on topics covered by the present guidelines which, to the knowledge of the team, have been recently published or are nearing completion:

- document by the European Council Steering Committee for Bioethics (CDBI) on the use of biological material of human origin for research and biobanks: "Recommendation of the Committee of Ministers to member states on research on biological material of human origin", adopted by the Committee of Ministers of the European Council on March 16th 2006. The document defines the ethical and juridical model with which member states' laws and guidelines regarding research on biological material and biobanks must comply;[(3)]

- document by the National Committee of Bio-security and Biotechnologies: "Guidelines for institution and certification of biobanks", issued on April 19th 2006; [4]
- document by the National Committee for Bioethics: "From pharmacogenetics to pharmacogenomics", dated April 21st 2006. The text illustrates the main issues in bioethics and describes how existing tools can be modulated to address them in the fields of pharmacogenetics and pharmacogenomics; [5]
- document by the National Committee of Bio-security and Biotechnologies working group on the genetic census of populations, dealing with different topics related to population studies; [6]
- document by the European Council Steering Committee for Bioethics (CDBI), on the application of genetic tests to the healthcare field viewed in a human rights perspective (Protocol to the Oviedo Convention, 1997) (awaiting completion);
- EMEA guidelines: "Concept paper on biobanks: pharmacogenetics and pharmacogenomics" [7] (awaiting completion). Will focus on collection, storage, future use, identification and regulatory requirements for samples collected during pharmacogenetic studies.

In conclusion, the potential of genetic research to have a profound impact on medicine and healthcare is now widely recognized and is the subject of guidelines like those indicated in 2004 by the U.K. Government in its Genetics White Paper. [8] By contrast, the basic requirements of the research which will make such changes possible have still mostly to be shared and harmonized.

Aim and scope of the guidelines

The Smith Kline Foundation and the Italian Society of Human Genetics promoted the drafting of the "Guidelines for Clinical Protocols in Genetic Research" as a contribution to the clear formulation of criteria for human genetic research. The aim of the present document is to provide guidance and advice for scientists contributing to clinical study design within genetic research, for members of the ethics committees to which the resulting research proposals are submitted, and for all those involved in genetic research applied to the improvement of human health.

The guidelines apply to any clinical study in genetic or genomic research, and to any study performed on human subjects in the course of which a biological sample is taken for genetic or genomic analysis.

The guidelines apply to any clinical study in genetic research, whether experimental (e.g. the administration of a new drug and the analysis of genetic polymorphisms on individual response) or observational (e.g. determination of the frequency of specific polymorphisms in a given population), irrespective of the type of research and study design. This is how the expression "genetic clinical research" is used in the present text.

It was decided that the present guidelines would address only issues specific to and characteristic of genetic clinical research, not clinical research as a whole. The rules, guidelines and principles that apply to clinical research at large are obviously a fundamental basis for a specific field such as genetic clinical research, but there is no shortage of published documents and data on these issues.

The present guidelines examine various topics related to genetic tests used for research purposes as opposed to clinical diagnostics. It is important to bear this distinction in mind when interpreting and applying the recommendations on the following pages, since the implications of genetic tests differ according to whether they are used for research or diagnostics (see the section "Genetic testing for diagnostic and research purposes").

There is no shortage of documents and data on genetic tests for diagnostic purposes. Of particular importance is the agreement drawn up by the Italian Health Ministry, the regional governments and the autonomous provincial authorities regarding the document "Guidelines for medical genetic activities", issued on 15th July 2004 by the Permanent Conference of the State, the Regions and the Autonomous Provinces of Trento and Bolzano (see Official Gazette n.224, 23rd September 2004) [9]. This document sets out to provide uniform national indications so that citizens can be guaranteed appropriate levels of healthcare, quality and standardization for the examinations and treatments concerned. Also important are the "Draft guidelines for quality assurance in molecular genetic testing", which the OECD recently issued for comment [10].

How the guidelines were drafted

Given the variety of fields involved in clinical genetic research and the range of potential issues, the contribution of experts in different disciplines was considered fundamental to an exhaustive discussion of the different topics concerned. The task of drafting the guidelines was entrusted to a multidisciplinary team of specialists in human genetics, bioethics, clinical medicine and clinical pharmacology, together with representatives of citizens, ethics committees, institutions, scientific societies and the pharmaceutical industry. The team was coordinated by the Italian Society of Human Genetics with the collaboration of the Smith Kline Foundation.

The team members gathered in plenary meetings in 2004 and 2005. During these meetings the main topics for correct implementation of genetic research were identified and discussed. Since the purpose of the document was to identify basic criteria and methods which could provide a reference basis for those involved in human genetic research, the spirit of the group discussions was that this aim could best be achieved by negotiating any differences in the positions of individual participants. Where opinions differed, the aim was to achieve the consensus of the group, even if this did not completely reflect the personal position of all participants. This made it possible to reach consensus on all fundamental topics, while personal opinions on various issues might of course differ from the stance expressed in the final document. The text in its present form, based on the group discussions, entailed a number of preliminary drafts. Each of these was made available to all the participants for their corrections and comments. For every topic, the document provides both highlights and detailed information so that different users' needs can be met. The document has thus been divided into two sections – the first contains a concise, functional statement of its content, while the second addresses topics in detail. Further insights are offered in a number of text-boxes. In addition, each chapter concludes with a section entitled "Summing and recommendations" which recapitulates the main themes. The final version of the document, approved by all participants, was issued on 23rd October 2006.

The experts and associations contributing to the drafting of the present document are listed in the "Authors" section on the opening pages.

Summary and Recommendations

The nature of the information generated from a genetic test differs according to the characteristics it investigates and its ability to predict the phenotype under analysis (e.g. a disease). This variability of genetic tests is reflected in their social and ethical implications. A fundamental distinction must be made between studies that investigate monogenic genetic traits (studied in protocols on rare genetic diseases, in which one gene is responsible for the disease) and multifactorial traits (studied in protocols on common diseases resulting from the interaction of multiple genes with each other and with the environment). The risk of discrimination and possible implications for relatives are lower in genetic tests on complex characters which, unlike those to identify monogenic characters, state only the estimated risk of a specific phenotype and not the certainty of its occurrence.

There is no reason to assume that (pharmaco)genetic information is qualitatively different from other medical information. The key to the implications of a medical test is the nature of the information it provides, and not whether it involves genetic data. Presymptomatic genetic testing (e.g. for Hungtington disease) is an exception, since the tests concerned are the only ones that can predict with certainty the future occurrence of a disease.

The term "genetic tests" identifies a variety of tests that can be classified according to their characteristics and purposes, this classification being reflected in their different ethical, social and legal implications. It is generally true that genetic tests for simple traits have greater ethical, social and legal implications than those for complex traits, and that genetic tests regarding diseases have greater ethical implications than those dealing with pharmacogenetics. The most important legal, ethical and social implications are probably those of presymptomatic testing. For such testing it can also be assumed that the information generated is qualitatively different from that obtained by other genetic and non genetic tests. The legal, ethical and social implications of genetic analysis regarding somatic mutations in cancer are probably not as great, though hereditary tumours will inevitably entail some reservations in this respect. An analysis of somatic mutations in cancer supplies information "restricted to the subject and the moment", concerning only the tumour already present. Such an analysis does not address the susceptibility of the patient or kin to this form of tumour, or the patient's risk of developing different types of tumour.

Diagnostic tests in a clinical setting have different ethical, social and legal implications from research testing. Only tests carried out for clinical purposes generate information about the health of the subject. Those performed for research purposes generally provide data that are not indicative of the the subject's health or genetic risk status, but are useful for scientific or statistical investigation.

The ethical, social and legal implications of genetic tests based on the analysis of single genetic factors also generally apply to genomic testing, which investigates many genes or extensive regions of the chromosome (including tests based on the analysis of RNA pattern variations, and thus of gene expression). They are also generally applicable to proteomics, or protein analysis.

The potential risk associated with genetic research is related to the information collected during the study or produced by genetic analysis. This risk also exists in non-genetic studies. The potential information-related risk associated with pharmacogenetic research is normally lower than for genetic testing in multifactorial disease, which is in turn lower than for genetic testing in monogenic disease.

Genetic research today normally benefits the scientific and medical community as well as society at large. Research contributes to scientific knowledge and its possible future clinical application. To date, genetic research has rarely resulted in direct benefits for the patient, but the first

pharmacogenetic studies are in progress and will generate information potentially useful for the health of the individual who has participated in the research.

All the rules and principles valid for clinical research are applicable to genetic research, but other specific issues must also be considered – e.g. management of biological samples (DNA, RNA).

Data protection for the subject participating in research is not only mandatory by law (Law 196/03 - Data Protection Act), but also an ethical duty. Where data handled for purposes of genetic research are non-anonymous, as defined in Art. 4 of the aforementioned Act, they are subject to these mandatory requirements. The data concerned may be either identifiable or coded, insofar as those using the biological material (e.g. the sponsor) can access the code and link the data to the subject. Data are not subject to the Data Protection Act when they are anonymized and cannot be linked to the identity of the individual; when they are anonymized and linkable to the subject's identity, they are exempt from the requirements of the Act only insofar as those using the biological material have no access to the code and thus to the subject's identity. For more detailed information on this topic, see the section ***"Level of identification".*** The investigator who handles non-anonymous data does not have to apply for the authorization of the National Data Protection Office, since a general dispensation (the official reference is to authorization 2/2005, 1.2 lett. *a*) applies in such cases; however, s/he must inform the National Data Protection Office that genetic data are being handled and obtain the subject's informed consent in this respect.

As a general criterion, as for any other biomedical research, every proposed genetic study must have a well founded rationale and clearly stated aims. When drafting a research protocol, it is sometimes not possible to describe the study rationale thoroughly and precisely, while the aims can be expressed only in general terms. This is often the case with the most recent research.

The study design must be appropriate for the achievement of its aims. In general terms, linkage studies in families are appropriate above all for study of Mendelian diseases or for initial identification of a genomic region related to the phenotype, but are limited by the need to include many or, if possible, very large families. Case-control studies are typically used in pharmacogenetic and epidemiological genetic research; possible confounds, such as stratification related to differences in ethnic group, age, sex, and other phenotype-related factors, must be taken into account.

Every genetic study, like any other biomedical research, must in principle involve a suitably large sample to achieve its aims. In practice, sample size in genetic research often depends on factors that are unknown at the beginning of the study.

It is useful to distinguish human tissue banks (biobanks) from collections of human biological samples created within a genetic study for the sole purposes of the study. The biobank differs from the simple collection of biological samples within a genetic study because it is created to fulfil more extensive purposes and, in particular, because one of the main goals of a biobank is to make biological samples available to those who ask for them. These differences have practical implications, for example in terms of how long samples are stored or how exactly permitted uses must be described.

The party promoting the research has the responsibility to guarantee that storage, analysis and transfer of samples will meet standards ensuring their integrity and safety; s/he must also guarantee that samples will be used for the purposes for which they were collected.

Samples and data collected or produced by the study can (or cannot) be linked, whether directly or indirectly, to the name of the subject who provided them. There are different scales to specify the levels of sample identification: five according to the EMEA document "Position paper on terminology in pharmacogenetics" [(11)], three in the document of the European Council Steering Committee for Bioethics entitled "Recommendation of the Committee of Ministers to member states on research on biological material of human origin" [(3)]. The most suitable level of sample identification must be chosen according to the protocol concerned. It must be borne in mind that samples and data which are anonymous or anonymized and non-linkable cannot be linked to the name of the subject who provided them. Optimum data protection is thus achieved, but at the same time the subject is not allowed to withdraw his/her consent, request destruction of the sample(s) concerned or receive information on the results of the analysis to which the material is subjected.

Storage and handling of samples after the conclusion of the study are appropriate and recommended if it is likely that stored samples will generate further scientifically useful information. In such cases, accurate information has to be provided, including storage methods, most likely use of the sample and the possibility that samples may be sent to other laboratories/research groups.

At any time and without limitation, the subject must be allowed to demand the destruction of the sample(s) supplied, unless anonymized and non-linkable.

The biological samples collected within a research study can also be used for future examinations included in the aims of the original study and described in the informed consent, without the need for fresh authorization and consent.

Samples can also be used also for a purpose not included among the aims of the original study provided that such use has been previously described in the protocol approved by the ethics committee and the patient has given specific consent. Otherwise, prior approval of the ethics committee and consent of the subject are needed (except for anonymized non-linkable samples).

The samples can be sent to other research groups or laboratories, on condition that this is done solely for purposes of scientific collaboration so that the sample can be subjected to certain procedures (for example, specific analyses); these can also be paid for, on condition that the guarantees regarding sample storage, possible uses and the level of data protection are exactly as specified in the informed consent. In such cases, this possibility must be mentioned in the informed consent and the required quality standards in handling and analyzing samples, data protection and compliance with uses specified in the informed consent must be guaranteed. The samples cannot be sold for a monetary consideration.

DNA, RNA and protein analysis technology is a rapidly developing field that continues to offer new opportunities. Genetic research quickly discovers new genes, polymorphisms or associations of known genes and polymorphisms with different phenotypes. The requirement for prior selection and description of genes, polymorphisms or analytical techniques in a genetic research protocol could limit investigators' opportunities to apply new techniques or explore new scientific hypotheses, thus reducing the likelihood of generating useful results. Conversely, this level of detail does not increase the level of protection for the patient. The protocol must detail the genes and polymorphisms analyzed only in extremely focused studies which set out to analyze the genes/polymorphisms concerned. Otherwise, indications must be given about the types of genes to be studied and methods of analysis.

In compliance with the principle stated in the European Council Convention on Human Rights and Biomedicine (art. 10), the Universal Declaration on the Human Genome (art. 5 *c*) and the UNESCO

International Declaration on Genetic Data (art. 10), and consistent with the provisions of the European Directive on Personal Data Protection, subjects who have participated in a research study have the right to receive their individual genetic results if they so wish, irrespective of whether such data are clinically useful. The informed consent must clarify the nature of such results and state whether they may be useful for the subject's health, so that the s/he has information on the basis of which to decide whether to ask for the results.

Genetic research might generate collateral information, i.e. information outside the study's objectives, or unexpected knowledge, i.e. information that was not expected to emerge from the study. This possibility is at present uncommon and usually the analysis reveals only what has been sought.

The European Commission document "25 recommendations on the ethical, legal and social implications of genetic testing" [(12)] makes a distinction between genetic counselling and genetic information and recommends that, with the exception of diagnostic and pre-symptomatic tests for severe diseases, the information generated by a genetic test can be disclosed to the patient by physicians who are not specialists in medical genetics but have received suitable training.

Individual results which have no immediate clinical utility must not be disclosed to any person other than upon request to the subject concerned, through the physician involved in the study. Individual results which may be useful for the health of the subject must be given to the study physician so that s/he can take them into due consideration and share them with the patient. The informed consent has to specify whether any results are expected to be useful for the subject's health and, if so, that the study physician will be duly informed of them.

As recommended in the Helsinki 2000 document on use of information by the study promoter (publication/non-publication of positive/negative results), and as for any other biomedical research, the global results of genetic research must be made public.

Patents on biotechnological inventions are based on the same requirements as any other invention: novelty, non-obviousness and utility. European Directive 98/44 (implemented in Italy as Law n.78 of 22nd February 2006) [(13)] also specifies that neither the human body nor the mere discovery of one of its parts is patentable, while a biological material isolated from its native environment or produced through a technical procedure, even if pre-existing in the natural state, is. This principle applies to the complete or partial sequence of a gene: the mere discovery of a gene sequence is not patentable, but the sequence becomes patentable if it has been isolated from the human body or produced through a technical procedure involving innovative processes, and if there is a description of a possible application for that sequence.

Regarding the possibility that the use of patented tests or sequences may lead to a situation of monopoly, a distinction must be made between patents and licenses. In the light of this distinction, one view of the matter is that monopoly situations arise as a result of licenses not being granted rather than of patents being registered.

The results of genetic research may create potential for commercial or patent exploitation. The subject who agrees to participate in the research must be aware of this possibility.

In human genetic research, the only insurable risk is physical and must be adequately covered by an insurance policy. It is currently not possible to insure against non-physical damage (any personal or moral damage resulting from the disclosure of information collected or produced by the study), such damage being difficult to estimate.

Correct planning and management of the investigation, in particular of samples and individual genetic data, help minimise the risk of any non-physical damage which might arise from genetic research. If the subject suffers non-physical damage not covered by the insurance policy, this can cause litigation to be settled in court or by the National Data Protection Office. In such cases, everything written in the protocol, informed consent, study archives or records of the laboratory where samples are stored and analyzed can have a considerable bearing on the issue of liability.

European Directive 2001/20 on Clinical Research, implemented in Italy as Law 211/2003, gives specific consideration to research on minors and mentally disordered subjects, providing the basic norms for genetic research on these categories. Since general caution is needed when performing genetic studies on minors and mentally disordered subjects, it is recommended that whenever possible consent be obtained from the subject concerned in addition to that of a parent/guardian. Studies on minors or mentally disordered subjects are legitimate when the disease studied is typical of the minor or mentally disordered subject, or when there is an advantage for the minor/mentally disordered individual, and even then only after studies have been carried out in adult subjects.

Both the protocol and informed consent for genetic research must be drafted and assessed in compliance with the same general criteria as for any biomedical research protocol. These documents must also contain detailed information regarding disclosure of individual results from the genetic study and the collection, storage and use of any biological samples involved. When the genetic investigation is simply part of a study with more extensive purposes it is appropriate that the subject should be free to choose whether to participate, without forfeiting the right to participate in the non-genetic part of the study.

It must be underlined that in some instances such a choice may not be possible. There are already ongoing pharmacogenetic studies in which patients are selected for different randomization groups according to their genotype. In such cases the genetic test is mandatory for participation in the clinical study and the patient cannot choose whether to accept or refuse it. Such studies may become more frequent in future, making it more difficult not only to distinguish clearly between clinical study and genetic study but also to give the patient a choice between agreeing to one or both of them.

A good model for the informed consent should achieve an optimum balance between the freedom of research and the thoroughness of the information provided to the individual. An indication of what should be contained in the protocol and informed consent of a genetic investigation is provided in the chapter **"Check lists for preparation and assessment of the protocol and informed consent for genetic research".**

The working group recommends that the following principles be borne in mind when drafting or assessing clinical research protocols in genetics.

1. **There is no reason to assume that information obtained from a genetic test is qualitatively different from other medical information, with the exception perhaps of presymptomatic genetic testing.**

2. **Protocols on rare diseases must be distinguished from those on complex diseases, because the two categories entail different specific issues.**

3. **Tests for research purposes, differing fundamentally from those for clinical purposes, entail more limited ethical, social and legal implications.**

4. **Pharmacogenetic tests do not have the same, or such great, implications as those for complex diseases, which in turn have lesser implications than those for simple traits.**

5. **The risk associated with genetic research is essentially related to the information c ollected during the research or produced by the genetic analysis. The magnitude of this risk depends on the type of research and the character studied.**

6. **The research must adopt proper measures for data protection and management of individual information in order to counter this risk.**

7. **At present, benefits deriving from genetic research are normally for the community, not the individual, but there are already ongoing studies which may benefit the subject and these could become more frequent in the near future.**

8. **Risks and benefits of the specific research must be adequately described in the informed consent.**

9. **Any genetic research must use the same rules and principles which regulate clinical research, and also deal appropriately with typical issues like those involved in management of DNA samples.**

10. **Individual data must be protected, in compliance with Law 196/03 and with a proper sense of professional ethics. Any research must comply with the requirements of Law 196/03 according to the type of data handled. Various additional measures are appropriate to help protect individual data. For instance, access to archives where data are managed (whether paper-based or electronic) and to rooms where biological samples are stored is strictly limited to persons involved in the research, while access to individual genetic results is limited to the patient and to the investigator in the case of results useful for the subject's health.**

11. **The study rationale and aims have always to be described. If it is not possible to describe them in detail, they must at least be stated in general terms.**

12. **Study design must be appropriate to the aims of the research. In general terms, linkage studies in families are mostly used to study Mendelian diseases or for initial identification of a genomic region related to the phenotype; case-control association studies are typically used for pharmacogenetic and epidemiological genetic research.**

13. **The protocol of the study must contain an explanation of the proposed sample size. If an accurate assessment of the necessary sample size is not possible, the reasons for considering the suggested sample size appropriate must be stated.**

14. **It is necessary to supply suitable information and guarantees about storage and possible uses of samples, particularly following the end of the study.**

15. **The level of identification of samples and data must be explicitly stated and must be suitable for the aims and the methods of the specific study.**

16. **It has to be guaranteed that the samples will be used solely for future uses described in the protocol and subject's informed consent.**

17. If the samples are to be put to uses not described in the protocol and informed consent, ethics committee approval and informed consent have to be obtained anew (except in the case of anonymized non-linkable samples).

18. Samples can be sent to other research groups or laboratories, on condition that this is done solely for purposes of scientific collaboration so that the sample can be subjected to certain procedures (for example, specific analyses); these can also be paid for, on condition that the guarantees regarding sample storage, possible uses and the level of data protection are exactly as specified in the informed consent. In such cases, this possibility must be mentioned in the informed consent and the required quality standards in handling and analyzing samples, data protection and limitation to uses specified in the informed consent must be guaranteed. The samples cannot be sold for a monetary consideration.

19. It is appropriate that the analysis of the collected samples should benefit from all the instruments made available by scientific and technological advances, to increase the likelihood of the research giving useful results. A research protocol rigidly restricted to the analysis of specific genes and polymorphisms might make it impossible to take advantage of such opportunities, while at the same time not providing greater safeguards for the patient. The protocol must detail the genes and polymorphisms analyzed only in extremely focused studies which set out to analyze the genes/polymorphisms concerned. For research with broader aims it is not necessary to specify in the protocol the genes and polymorphisms to be studied and the analysis techniques; indications about the types of genes to be studied and methods of analysis must nevertheless be given. It is appropriate to mention some examples if available.

20. Subjects who have participated in a research study have the right to receive their individual genetic results if they so wish, irrespective of whether such data are clinically useful. The informed consent must clarify the nature of such results and state whether they may be useful for the subject's health, so that the s/he has information on the basis of which to decide whether to ask for the results.

21. Individual results which have no immediate clinical utility must not be disclosed to any person other than upon request to the subject concerned, through the physician involved in the study. Individual results which may be useful for the health of the subject must be given to the study physician so that s/he can take them into due consideration and share them with the patient. The information sheet has to specify whether any results are expected to be useful for the subject's health and, if so, that the study physician will be duly informed of them.

22. It is rare for genetic research to produce unexpected knowledge or collateral information. This issue has to be managed only when this information is important for the health of the subject. It is neither practicable nor appropriate that the informed consent should include the choice as to whether any unexpected knowledge or collateral information should be made known to the subject, since the rarity and unpredictability of such events does not allow correct prior information enabling the patient to make a fully informed choice.

23. The production and dissemination of the results, as for any clinical research, is a duty required by the Helsinki 2000 document on use of information by the study promoter

and conistent with the subject's wish to contribute to the advancement of scientific knowledge by participating in the research.

24. Investigators who participate in the research should receive a report containing the overall results.

25. Samples collected within a study cannot be sold or bought for profit. However, the research results may create commercial advantages and/or patents for the research promoter. When a patent or commercial exploitation is envisageable, this must be stated in the informed consent and it must be clarified that there will be no economic benefit for the subject.

26. An insurance policy is essential to cover risk of material damage. To protect the subject (but also the investigator and the study promoter) from non-physical risks, it is important that the protocol and informed consent clearly state critical issues regarding non-material risks and related liabilities. The most critical issues are storage, use and level of identification of samples and data, together with the disclosure of individual genetic analysis results and correct, exhaustive information to the patient about possible risks.

27. Studies on minors or mentally disordered subjects are legitimate when the disease studied is typical of the minor or mentally disordered subject, or when there is an advantage for the minor/mentally disordered individual, and even then only after studies have been carried out in adult subjects. Whenever possible consent must be obtained from the subject concerned in addition to that of a parent/guardian.

28. The protocol and informed consent for genetic research must be drafted and assessed in compliance with the same general criteria as for any biomedical research protocol. These documents must also contain detailed information regarding disclosure of individual results from the genetic study and the collection, storage and use of any biological samples involved.

29. When the genetic investigation is part of a clinical trial, specific informed consent must be obtained regarding the subject's involvement in the genetic testing. The subject should be allowed, if s/he so desires, to participate in the clinical study but not the genetic testing. It must be underlined that in some instances such a choice might not be possible. There are already ongoing pharmacogenetic studies in which patients are selected for different randomization groups according to their genotype. In such cases the genetic test is mandatory for participation in the clinical study and the patient cannot choose whether to accept or refuse it. Such studies may become more frequent in future, making it more difficult not only to distinguish clearly between clinical study and genetic study but also to give the patient a choice between agreeing to one or both of them.

30. The completeness and comprehensibility of information given to the subject are particularly important for genetic research.

31. The following chapter, "Check-list for drafting and assessment of the protocol and informed consent in genetic research", indicates what must be included in them.

Check-list for drafting and assessment of a genetic research protocol and informed consent

As for any other biomedical research protocol, the ethics committee must assess the scientific correctness of the genetic research and its ethical justification. This section presents check-lists of the contents and assessment criteria of the protocol and informated consent in human genetic research. Only those items which are critical or peculiar to genetic research are mentioned. In addition to these specific features, items to be assessed in all biomedical research must also be considered. The check-lists, based on the working group's experience, should not be taken as an absolute standard for the contents and assessment methods of the protocol and informed consent. For correct application of the check-lists, we recommend that the reader consult the specific sections of the present guidelines which address individual issues, as indicated in the right-hand column of the table.

Check-list for the drafting and assessment of a human genetic research protocol

Contents	What to assess	Relevant section of the guidelines
Research rationale	Is the research rationale stated?	Study rationale and aims page 46
Aims of the study	Are the aims of the study sufficiently clarified? Are expected future scientific, clinical and social implications mentioned (if relevant)?	Study rationale and aims page 46
Study design	Is the study design correct and appropriate?	Study design page 48
Populations: size and characteristics	Are the size and characteristics of the population appropriate?	Sample size page 50
Information on scheduled genetic analyses	Are the scheduled genetic analyses made sufficiently clear? If not, have the reasons been explained?	Genetic analysis: which level of information page 64
Biological samples: collection, use, storage	Are the following specified? : • collection methods • level of identification and implications for the subject • time of storage • uses to which samples can be put • what will happen to samples on expiry of their scheduled storage period • guarantee of safe storage • responsibility for storage • possibility for the subject to ask for the destruction of the sample (except for anonymized non-linkable samples) • whether samples can be sent	Biological samples: level of identification, storage, uses page 52

	to other laboratories and, if so, guarantee that the level of subject and sample protection stated in the protocol and informed consent will be maintained.	
Confidentiality of information	Is there an adequate description of the methods of data protection? In the case of non-anonymous data, is it specified who will be responsible for data handling at the different stages of the study? Is it specified which categories of subjects will be authorized to access data?	Data protection page 41
Methods for withdrawal of subject's consent	Is the subject's right to withdraw from the study specified and, if so, is there an indication of what will happen to the samples and information generated? In the case of anonymized non-linkable samples, are the implications regarding possibility of withdrawal from the study and sample destruction explained?	Data protection page 41 *Obtaining the subject's informed consent* *page 42* *Level of identification* *page 56*
Study results: individual and overall results	Is it specified whether the study results may be useful for the health of the subject? Is it stated that subjects can demand access to their individual results (even if not useful for their health, except where samples are anonymized and non-linkable)? Is it stated to which other persons the subject's individual results can be disclosed (e.g. study physician, relatives, etc)? Is the dissemination of individual genetic results consistent with the aim of ensuring both that the subject's health is safeguarded and that data are protected? Is it stated whether the overall study results will be disclosed to the investigator?	Study results and accessibility page 66 *Benefits of human genetic research* *page 39*

Check-list for the drafting and assessment of the informed consent for human genetic research

Contents	**What to assess**	**Relevant section of the guidelines**
Study aims	Are the study aims adequately explained and understandable?	Study rationale and aims page 46
Voluntary character of enrolment	Is it clear that the subject has the right not to take part in the study or to withdraw from it? In the case of anonymized non-linkable samples, are the implications regarding the possibility of withdrawal from the study clear?	Protocol and informed consent page 80 *Level of identification* *page 56*
Implications of enrolment	Are the implications of enrolment for the subject specified? Is the scientific/moral value of participation mentioned? In studies involving other family members, is this possibility stated together with its possible implications for them?	Risks and benefits of human genetic research page 35 Genetic tests and ethical, social and legal implications page 31
Biological samples: collection, uses, storage	Are sample collection methods defined? Is what will be done with the provided sample described comprehensibly? Is it guaranteed that the sample will be put only to the uses specified in the informed consent? Are the level of identification of the sample (i.e. how directly the sample will be linkable to the subject) and the implications of this for the subject described? Are the duration of sample storage and what will happen to the sample after such time specified? Are there guarantees of safe storage? Is the subject's right to ask for destruction of the sample(s) s/he provided clear? In the case of anonymized non-linkable samples, are the implications for sample destruction explained? Is it specified whether the sample may be sent to other laboratories? If so, is the level of subject and sample protection stated in the informed consent guaranteed?	Biological samples: level of identification, storage, uses page 52 *Level of identification* *page 56*
Possible risks	Are the possible risks arising from	Risks and benefits of human

	disclosure of genetic information and the measures for protecting the confidentiality of such information described? Are other possible risks for the subject or for kin stated (if relevant)?	genetic research page 35
Direct advantages for the subject or the community	Is it clarified whether direct advantages can be expected for the subject? Are possible advantages for the community described?	Risks and benefits of human genetic research page 35
Confidentiality of information	Is it stated who will be able to trace the subject's identity starting from the individual data concerned, and in which situations? Is it stated which people will be able to access the subject's data? Is it stated to whom, besides the subject, individual results will be disclosed? In the case of non-anonymous data, is it stated who is responsible for data handling at the different stages of the research?	Data protection page 41
Methods for withdrawal of informed consent	Is the subject's right to withdraw from the study specified and, if so, is there an indication of what will happen to the samples and information generated? In the case of anonymized non-linkable samples, are the implications for the possibility of withdrawal from the study and sample destruction explained?	Data protection page 41 *Obtaining the subject's informed consent* *page 42* *Level of identification* *page 56*
Study results: individual and overall results	Is it specified whether the study results may be useful for the health of the subject? Is it stated that subjects can demand access to their individual results (even if not useful for their health, except where samples are anonymized and non-linkable)? Is it stated to which other persons the subject's individual results can be disclosed (e.g. study physician, relatives, etc)? Is the dissemination of individual genetic results consistent with the aim of ensuring both that the	Study results and accessibility page 66

	subject’s health is safeguarded and that data are protected?	
Commercial and patent uses	Is the possibility of commercial or patent use of the results mentioned, specifying any rights of the subject in this respect?	Commercial and patent rights page 70

Is genetic information different from other types of medical information?

Human studies involving genetic tests generate genetic information on the participating subjects. Before venturing any considerations on the rules appropriate to clinical protocols in human genetics, the nature of genetic information and the extent of its implications must be understood. The nature of the information generated from a genetic test depends on the trait investigated and the ability of the test to predict the analyzed phenotype, for example whether the individual will develop the disease concerned. The resulting social and ethical implications will vary accordingly.

Penetrance and expression of the gene

A fundamental distinction must be made between studies that investigate monogenic genetic traits and multifactorial traits. Monogenic or simple traits are those studied in protocols regarding rare genetic diseases, in which one gene is responsible for the disease. Multifactorial or complex traits are those studied in protocols regarding common diseases which result from the complex interaction of multiple genes with each other and with the environment. (See the text box ***"Distinguishing monogenic and multifactorial genetic traits"*** and **Table 5:** *Differences between simple and complex traits*).

Distinguishing monogenic and multifactorial genetic traits

A genetic trait is monogenic (or simple) when a single gene is responsible for the observed phenomenon (phenotype), for example the disease. Thousands of monogenic diseases have been identified, and they usually have frequencies lower than 1/1000 and involve no more than 2-3% of the entire number of individuals suffering from diseases. They are also called rare, Mendelian or hereditary genetic diseases. Their fundamental characteristic is that they are caused by a single gene, called the disease-gene. At the time of writing (14th September 2006), the "Online Mendelian Inheritance In Man" site [(14)] indicates:

- 10981 genes with known sequences
- 383 genes with known sequence and phenotype
- 1969 phenotypes with a known molecular mechanism
- 1548 Mendelian phenotypes or loci with an unknown molecular mechanism.

From the diagnostic point of view, identifying a mutation in a gene which causes a disease makes it possible to diagnose a genetic disease which may be already present or will surely develop in future. This disease is mostly congenital, i.e. present at birth, and genetic diagnosis in a clinical setting makes it possible to confirm or better define a diagnosis based on clinical criteria.

Far more rarely, the disease is not present at birth and develops in adult life. In such cases, as in Huntington disease, a genetic test can accurately predict that the individual will develop the disease later in life. In the great majority of cases there are currently no adequate therapies for monogenic diseases. These are diseases that, once diagnosed, take a relatively predictable course, even if there is a considerable variety of different mutations in the same disease – sometimes related to clinical manifestations, sometimes not. Another fundamental point is that these diseases are inherited according to the rules codified by Mendel and follow a predictable pattern of transmission. There is thus a direct connection between the presence of the gene and its direct expression on the phenotype, which could possibly lead to discrimination against the individual concerned and blood relatives.

A genetic trait is multifactorial (or complex) when the observed phenotype, a disease or a drug response, results from the effect of multiple genes and their interaction with the environment. By environment is meant any non-genetic factor, such as diet, lifestyle, age, gender, concomitant diseases or use of several drugs at the same time.

Examples of multifactorial diseases include cardiovascular and psychiatric disorders, asthma and other respiratory diseases, diabetes and cancer. These diseases account for the vast majority of observed morbidity, involving one person out of two in the western world. They are called multifactorial, complex or common diseases. Genetic influence on individual drug response is with few exceptions a multifactorial trait, this being a topic of investigation in pharmacogenetic research. The complexity of these traits does not make it possible to predict with certainty the occurrence or development of a disease or drug response. The genetic test will allow only a determination of susceptibility, in other words identification of greater or lesser genetic risk. It is this genetic susceptibility, and not the disease, which can be transmitted following an unpredictable, non-Mendelian pattern, because it will be modified by other genetic and environmental factors in descendants. Recent data indicate that genetic influence in rare diseases is much more varied than previously thought, and that the distinction between monogenic and multifactorial traits is less categorical than previously supposed. It has been ascertained that in monogenic diseases, genes other than that determining the disease are involved, affecting its course and severity. These are known as modifier genes. Environmental factors too have been shown to play a role, albeit a small one. In addition, a protein encoded by a specific gene can perform multiple functions and thus influence several aspects of the phenotype at the same time. These and other observations have made it necessary to revise previous views of genetic traits and diseases. All diseases are actually complex, caused by genes and the environment, monogenic diseases being those in which one gene with a clearly prevalent effect in relation to other genes and the environment is alone in determining the disease.
It is nevertheless on the whole still true, as already indicated, that monogenic traits exert a direct action and multifactorial traits an indirect action on the phenotype.

Table 5: *Differences between simple and complex traits*

	Simple	**Complex**
Inheritance	Monogenic	Multigenic
Environment	No	Yes
Mutations	Causative	Predisposing factors
Distribution	Bimodal	Continuous
Number	Thousands	Dozens
Frequency	Rare	Common

The ability of a genetic test to predict with certainty the studied phenotype, for example whether the subject will develop the disease, depends on the *penetrance* of the analyzed mutation. Only tests for simple traits (for example the CFTR gene test for cystic fibrosis) have high predictive value, while those for complex traits have low predictive value. A genetic test for a simple trait, with high penetrance, allows a reliable phenotype diagnosis and provides information on the mechanism of trait inheritance. Conversely, a genetic test for a complex trait, with low penetrance (for example the MTHFR gene test for ischemic cardiopathy, deep venous thrombosis and stroke) allows only an estimate of risk variation for phenotype occurrence and does not make it possible to identify the mechanism of phenotype inheritance. (See the text box ***"The effect of genes on phenotype: penetrance"*** and **Table 6:** *Penetrance and the effect of genes on phenotype*).

If genetic research includes a test for a complex trait, such as those related to common cardiovascular, psychiatric, respiratory, metabolic or oncological diseases or a pharmacogenetic test, the test result will in no way make it possible to identify with certainty those individuals who will develop the disease (or will have an altered drug response), but may possibly identify subjects with a modified risk of disease or of altered drug response.

A genetic test for a simple trait gives a direct assessment of a particular phenotypic effect, which may be a basis for *genetic discrimination*. The risk of discrimination with tests for complex traits is much lower because there is no direct assessment of correlation with a certain phenotype and the risk of developing it is merely estimated.
Equally, in this kind of test it is inappropriate to envisage a risk that the information obtained from the DNA of a person may affect his family too, because the pattern of inheritance is complex and unpredictable.

The effect of genes on phenotype: penetrance

The term "penetrance" indicates the frequency with which a given genotype will actually result in the corresponding phenotype (for example a disease). Simple traits have high penetrance, while complex traits have low penetrance. High penetrance indicates the existence of a major gene that determines the phenotype, while low penetrance indicates the effect of several genes, none of them more important than the others. Penetrance determines the possibility to predict with varying degrees of accuracy the phenotype associated with the mutation. Only genetic tests based on analysis of mutations associated with high penetrance phenotypes allow sure prediction of the phenotype.

Table 6: *Penetrance and the effect of genes on phenotype*

Penetrance	Phenotype	Examples of genetic diseases	Pharmacogenetic examples
High (one major gene)	Certain	Cystic fibrosis	Hypersensitivity to succinylcholine
Low (no major gene, multifactorial)	Probable	Venous thrombosis, Factor V Leiden	Impaired response to salbutamol

Since much of the genetic knowledge currently available has been developed on simple traits, the characteristics and implications of Mendelian traits are often incorrectly transferred to complex traits. For this reason, the predictive power of genetic tests for complex traits and their potential to modify clinical practice are sometimes overestimated, as are the possible ethical implications entailed.
The working group considers that part of the ongoing debate on genetic tests and clinical protocols in human genetics suffers from an excessive focus on the features of tests for simple diseases; different, more specific considerations are necessary for complex genetic traits, which most clinical protocols in human genetics are are concerned with.

"Genetic exceptionalism" or common medical information?

There has been much debate on the comparison between genetic information and common medical information. (See the text-box "***The debate on the nature of genetic information***").

The position that prevails today is that genetic information is non-exceptional, a stance taken by various prestigious institutions. In the document "Pharmacogenetics: Ethical Issues"[15], the Nuffield Council on Bioethics states the following: "There is no reason to assume that genetic information, including pharmacogenetic information, is qualitatively different from other medical information. The nature of the information provided by a medical test is the key to considering its implications, not whether the test involves genetic data".
A similar position is stated in the document of the National Committee for Bioethics "From Pharmacogenetics to Pharmacogenomics"[5], and in recommendation n. 3 of the document "25 recommendations concerning the ethical, legal and social implications of genetic tests" [12] – though the latter acknowledges that the perception of exceptionalism is widespread: "*It is recommended to avoid genetic exceptionalism internationally, in the UE and in the single Member States. However*

the widespread perception of the diversity of genetic tests should be recognized and considered".
The most recent report of the CIOMS (Council for International Organization of Medical Sciences) is entitled "Pharmacogenetics - towards improving treatment with medicines" [(16)]. This document states, in the chapter on "Genetic testing, genetic data and genetic information", that: "the distinction between genetic and non-genetic tests lacks a scientific rationale and is not helpful as a basis for indications on how to use the information or on how to protect the information and the individual concerned from bad use of such information".
The topic of genetic exceptionalism has been discussed in depth in an article published in the Hastings Centre Report of July-August 2005 [(17)]. The author points out that most scientific opinion is against the hypothesis of genetic exceptionalism, while legislators are "fascinated" by it. After examining pros and cons, the author concludes that genetic exceptionalism generates laws and social policy of little value, and that special laws for genetics reinforce the stigma that surrounds people with a genetic disease, since they are treated differently from people with non-genetic diseases.

The debate about the nature of genetic information

The debate about whether genetic information differs from common medical information involves two principal positions.
Those who maintain it does not argue that most of the information collected and the data produced during normal medical practice are more predictive of illness and risk than any genetic information. For instance, blood cholesterol levels are more predictive of cardiovascular risk than a genetic test, and smoking is certainly much more predictive of a range of risks than a genetic test. This means that medical data would provide a simpler basis than genetic data for calculating a hypothetical health insurance premium related to risk levels.
Those who consider genetic information exceptional argue that it is permanent and may affect kin.
The issue is actually even more complex and the distinction between the two positions far from clear-cut. The fact that genetic information is permanent does not in itself entail an inevitable effect on the risk of disease. An individual who is informed that s/he has an increased risk of cardiovascular disease because of genetic predisposition can adopt an appropriate life-style (for instance by giving up smoking, starting to exercise or going on a correct diet) and, though this will not affect genetic risk, overall cardiovascular risk will be reduced.

Having examined the arguments in favour of the exceptionalism and non-exceptionalism of genetic information, the working group considers that there is more support for the view that such information is not exceptional, with the exception of presymptomatic genetic tests. These are the only tests that can predict with certainty the future occurrence of a disease. To date, they are available for only a few dozen genetic diseases and involve about 1% of all patients. (See **Table 7**: *Categories of medical genetic test"* and the chapter on **"Genetic tests and ethical, social and legal implications"**).
The arguments for and against the exceptional nature of genetic information are summed up in the following table.

Genetic "exceptionalism": arguments for and against

FOR	AGAINST
Data produced by a genetic test are invariable: since a genetic alteration and any resulting stigmatization are present throughout an individual's life, knowledge of such information has a very strong impact	The same consideration applies to non-genetic diagnosis of any incurable chronic disease – for example, diagnosis of Alzheimer's disease identifies a non-reversible condition. Here too, knowledge of this "datum" has a lifelong effect and can result in stigmatization

Data produced by a genetic test may have serious emotional ("morbidification"), social and economic implications.	The same risk may occur even if genetic tests are not performed: the mere identification of several individuals in a family with a certain disease leads to suspicion of a family predisposition and possible "morbidification". Any diagnostic tests, even if non-genetic, can have a social and economic impact.
Information is potentially misleading, incorrect and in some cases inconclusive (concept of probability and risk); there is rarely any possibility of treatment.	This is true also for many non-genetic tests: cholesterol levels are not always predictive of cardiovascular risk and PSA is not always predictive of prostate cancer risk. The same applies to possibilities of treatment.
A genetic test on an individual can provide important knowledge for other members of the family and vice-versa (concept of the biological group – see the EC text "Working document on Genetic Data", Working Party on the Protection of Individuals, March 2004).	The same can be said of non-genetic diagnosis. Doctors have for many years used family history, as have insurance companies for calculation of premiums. For instance, early cardiovascular events in several members of a family imply increased risk for other family members, even without genetic testing.

The information produced by a genetic test tends to be perceived as inevitable. The working group believes that the reason for this perception is not an intrinsic difference between genetic and non-genetic information, but the absence of effective treatments for the great majority of Mendelian genetic diseases. The advent of gene therapy in clinical practice would end this sense of inevitability in the same way as the discovery of antibiotics and vaccines did for infectious diseases, long considered impossible to prevent or treat. Even without gene therapy, examples such as phenylketonuria indicate that, just as as for many other diseases, the real problem is more the current lack of appropriate therapeutic tools. The emotional impact of diagnosis, the media and contingent circumstances should also be taken into account, reinforcing as they do the idea of inevitability often associated with genetic diseases.

Summing up and recommendations

The nature of the information generated from a genetic test depends on the trait investigated and the ability of the test to predict the analyzed phenotype, for example whether the individual will develop the disease concerned. The resulting social and ethical implications will vary accordingly.
A fundamental distinction must be made between studies that investigate monogenic genetic traits and multifactorial traits. Monogenic or simple traits are those studied in protocols regarding rare genetic diseases, in which one gene is responsible for the disease. Multifactorial or complex traits are those studied in protocols regarding common diseases which result from the complex interaction of multiple genes with each other and with the environment.
The risk of discrimination with tests for complex traits is much lower than with tests for simple traits, because there is no direct assessment of correlation with a certain phenotype and the risk of developing it is merely estimated.

There is no reason to assume that genetic information, including pharmacogenetic information, is qualitatively different from other medical information. The nature of the information provided by a medical test is the key to considering its implications, not whether the test involves genetic data. The nature of the information provided by a medical test is the key to considering its implications, not whether the test involves genetic data. Presymptomatic genetic tests (e.g. for Hungtington

disease) are an exception, being the only tests that can predict with certainty the future occurrence of a disease.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

• there is no reason to assume that information from a genetic test is qualitatively different from other medical information, with the exception perhaps of presymptomatic genetic testing

• rare disease protocols must be distinguished from complex disease protocols because the two involve different specific issues.

Genetic tests and ethical, social and legal implications

Categories and development of genetic tests for medical purposes

The term "genetic tests" defines a varied group of tests of which almost the only common feature is that they are based on the analysis of DNA. They are classifiable according to their different characteristics and aims (see **Table 7**: *Categories of medical genetic tests*).

Table 7: *Categories of medical genetic tests*

Monogenic traits	**Complex traits**
DIAGNOSTIC (SYMPTOMATIC) On affected individuals To establish or confirm a diagnosis PRESYMPTOMATIC (PRECLINICAL) On healthy individuals If the test is positive, the disease is certain to occur HETEROZYGOTE IDENTIFICATION For autosomal recessive diseases	PREDICTIVE (OR SUSCEPTIBILITY TEST) For complex diseases or to predict drug response. Identify predisposition or resistance to a disease or to the effects of a drug. Probabilistic.
DISEASE DIAGNOSIS	**DETERMINATION OF SUSCEPTIBILITY**

With advances in scientific knowledge of complex genetic characters and how they affect the phenotype, genetic tests are increasingly entering the domain of molecular medicine and no longer of traditional medical investigation. Genetic tests will in future be more and more concerned with complex rather than simple traits, their implications shifting more towards the individual than the family and extending increasingly from prenatal diagnosis to adult life; risk determination and prediction will complement the familiar diagnostic and preventive use of genetic testing, while the methodology will move from chromosomes and genes towards genome analysis (see **Table 8**: *Development of genetic tests for medical purposes*). Clinical protocols in genetic research will be increasingly focused on complex diseases. Investigations in the complex disease field aim at defining the effects of single genes in complex phenotypes and identifying genotypes with a major effect. Such research should create the scientific basis for routine genotyping, as a way of determining hereditary predisposition to common diseases (predictive medicine) and as a starting point for genotype-specific therapy (pharmacogenetics).
There is still some disagreement about the definition of a genetic test as a test based on the analysis of DNA sequence variations, it being argued that the definition should be extended to anything influenced by genetics – thus including genomic and proteomic analyses.

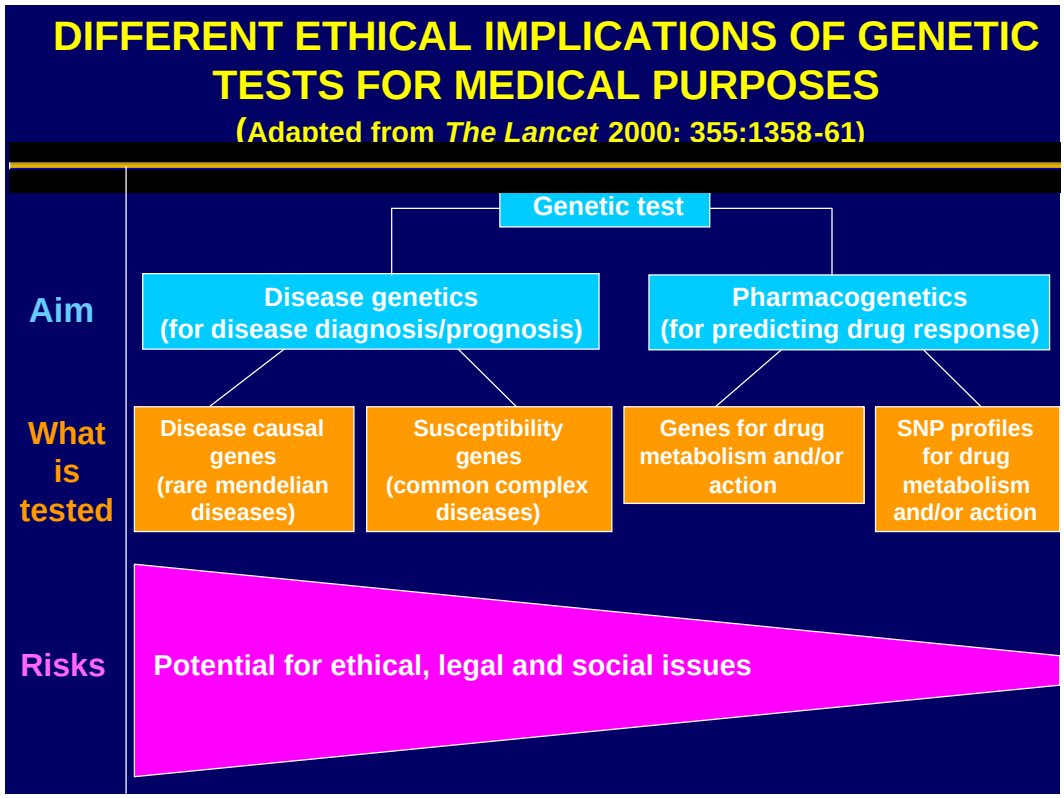
Table 8: *Development of genetic tests for medical purposes*

	Traditional Medicine	Molecular Medicine
Phenotype	Simple	Complex
Subject	Family	Individual
Method	Chromosome-/gene-based	Genomic
Timing	Prenatal	Adult life
Purpose	Prevention-diagnosis	Determination of susceptibility

Ethical, social and legal implications of the various kinds of genetic test

The working group considers that, although the ethical, social and legal implications of every genetic test can be correctly examined only on a case by case basis, it is generally true that genetic tests for simple traits have greater ethical, social and legal implications than those for complex traits, and that genetic tests regarding diseases have greater ethical implications than those dealing with pharmacogenetics. The working group considers that the diagram showing the ethical, social and legal implications of genetic tests (see Figure 1), published in *The Lancet* in 2000[18] as part of the article "Pharmacogenetics and future drug development and delivery", provides a valid general reference framework for discussing the risks of genetic testing within genetic clinical protocols.

Figure 1: *Different ethical implications of genetic tests for medical purposes*



Presymptomatic genetic testing and genetic analysis of somatic mutation in cancer should be considered as a specific topic in their own right. Presymptomatic genetic testing is the only genetic testing that can predict with certainty the future occurrence of a disease (see **Table 7**: ***Categories of medical genetic tests***). A 20 year-old subject with a positive test for Hungtington disease will certainly develop this illness later in life. It is thus readily understandable that the form of investigation with the greatest legal, ethical and social implications is probably presymptomatic testing. For such testing, it can also be assumed that the information generated is qualitatively different from that obtained by other genetic and non genetic tests. Presymptomatic genetic testing is available for a few dozen hereditary diseases, including Hungtington disease, myotonic dystrophy, and adult polycystic kidney disease, accounting for about 1% of patients.

Genetic analysis of somatic mutation in cancer is another area worthy of special consideration, albeit with some reservations for hereditary tumours. DNA extracted from tumour tissue is different from that of the host subject. Analysis of somatic mutations in cancer gives information "restricted to the subject and the moment", concerning only the existing tumour and not the individual's predisposition to the same or other tumours, or the likelihood of any other family members having an increased risk for that tumour. This kind of analysis thus entails less marked legal, ethical and social implications.

The above considerations regarding ethical, social and legal implications for *genetic tests*, based on the analysis of DNA sequence variations, can in general terms be extended to genomic tests, including those based on the analysis of RNA pattern variations and therefore of gene expression. They also apply generally to proteomics, which deals with protein analysis. The common denominator is the predictive capacity of the information generated and the implications of this, irrespective of the methods used.

It is worth repeating that the nature of the information produced by any medical test is the key to considering its implications, not the methodology used.

Genetic testing for diagnostic and research purposes

The working group believes that, as for non-genetic tests, it is important to distinguish genetic testing with diagnostic goals in a clinical setting from testing for research purposes, such as that performed in a genetic study. This distinction must always be borne in mind in any considerations regarding clinical research protocols in human genetics.

If a genetic test is performed with a diagnostic purpose, it means that the genotype-phenotype correlation on which the test is based is known and reproducible, and that the test produces information which is considered clinically useful. A genetic test performed in a research context has the purpose of identifying a genotype-phenotype correlation and as a rule produces information which is not in itself immediately useful in the clinical setting, but for statistical and scientific research. The potential implications for the subject and kin differ greatly, as does the likelihood of discrimination or other disadvantages for the subject: a genetic test for clinical purposes produces information on the health of the subject, while research testing does not. For instance, a positive genetic test for diagnosis of hereditary breast cancer identifies increased risk of developing the disease in the person tested. A genetic test for hereditary breast cancer in a research context does not aim to establish a diagnosis; from the overall data produced by the study it might prove possible to identify a genotype-phenotype correlation, which in itself does not usually make it possible to establish individual test subjects' risk of developing the disease. Further confirmatory and validation studies are necessary to create an appropriate basis for diagnostic use of the test.

The diagnostic use of genetic tests is regulated in Italy by the "Guidelines for medical genetics activities", issued by the Permanent Conference of the State, the Regions and the Autonomous Provinces of Trento and Bolzano following the Agreement of the Health Ministry, the Regions, the Autonomous Provinces of Trento and Bolzano[(9)]. These Guidelines were drafted to give uniform

national indications so that citizens can be guaranteed appropriate levels of healthcare, quality and standardization for the examinations and treatments concerned.
By contrast, the aim of the present guidelines is to provide recommendations on genetic tests for research purposes.

Summing up and recommendations

The term "genetic tests" defines a varied group of tests of which almost the only common feature is that they are based on the analysis of DNA. Classifiable according to their different characteristics and aims, they vary in their ethical, social and legal implications.

It is generally true that genetic tests for simple traits have greater ethical, social and legal implications than those for complex traits, and that genetic tests regarding diseases have greater ethical implications than those dealing with pharmacogenetics.

The form of investigation with the greatest legal, ethical and social implications is probably presymptomatic testing, the only form of genetic testing that can predict with certainty and the future occurrence of a disease. For such testing, it can also be assumed that the information generated is qualitatively different from that obtained by other genetic and non-genetic tests.

Genetic analysis of somatic mutation in cancer entails less marked legal, ethical and social implications, albeit with some reservations for hereditary tumours. DNA extracted from tumour tissue is different from that of the host subject. Analysis of somatic mutations in cancer gives information "restricted to the subject and the moment", concerning only the existing tumour and not the individual's predisposition to the same or other tumours, or the likelihood of any other family members having an increased risk for that tumour.

Testing for diagnostic purposes in a clinical setting has different ethical, social and legal implications from testing with research aims. Only clinical testing generates information on the individual's state of health. A genetic test performed in a research context as a rule produces information which is not in itself indicative of the individual's state of health, but is useful for statistical and scientific research.

The above considerations regarding ethical, social and legal implications for *genetic tests*, based on the analysis of single genetic factors, can in general terms be extended to genomic tests, concerned with large numbers of genes or extensive chromosomal regions (including testing based on analysis of RNA pattern variations and therefore of gene expression). The same also applies generally to *proteomics*, which deals with protein analysis.

In the drafting or assessment of clinical research protocols in the genetic field, the working group recommends that the following points be given due consideration:

- **there are essential differences between tests for research purposes and those for clinical aims, the former having less marked ethical, social and legal implications**
- **pharmacogenetic tests have lesser implications than those for complex diseases, which in turn have lesser implications than those dealing with simple traits.**

Risks and benefits in human genetic research

"Non-physical" risk

The potential risks of genetic research can be divided into the two categories of "physical", typical of any research, and "non-physical", related to genetic information produced by the study. See **Table 9**: *The potential risks of genetic research* and the text-box "***Stigmatization and discrimination***".

Table 9: *The potential risks of human genetic research*

Physical risk	**Non-physical risk**
• Sample collection	• Psychological damage to the subject due to awareness of individual results • Damage to the subject's privacy due to the disclosure of individual results • Damage caused by stigmatization of the subject's ethnic group
PHYSICAL RISK related to the sample collection method, or to drugs administered during the study	**"INFORMATION-RELATED" RISK** related to disclosure of information

Physical risk in genetic research originates from the procedure of sample collection or from any drugs administered during the study. In many cases this risk is limited or non-existent – for example, if the research involves only collection of a blood sample for DNA extraction.
There is a second type of risk, which can be called "non-physical", involving:

• psychological damage to the subject through awareness of individual results
• damage to the individual and breach of confidentiality caused by disclosure of individual results to third parties
• damage caused by stigmatization of the subject's ethnic group, if the research involves or focuses on a specific ethnic group.

Non-physical risk, which we could call "information-related" risk because it is related to information collected during the research or produced by genetic analysis, is not peculiar to genetic research – much of the information and data from a normal clinical study is more predictive than any genetic information (levels of cholesterol, smoking, etc.).
However, these risks are related to the disclosure of information and can thus be reduced or removed by adequate data protection and appropriate management of individual data produced by the research.
Recognizing that there are no scientific or medical grounds for considering people with specific genotypes or belonging to specific ethnic groups as "genetically disadvantaged", the working group believes that the wrong and sometimes unintentional creation of population groups catalogued or perceived as "genetically disadvantaged" must be opposed, and that research, particularly if public, should target medicines for patients with less frequent genotypes which would not be of interest for private research. For an in-depth discussion, please refer to the following text-box, ***"Stigmatization and discrimination".***

Stigmatization and discrimination

The development of genetic research has been accompanied by concern lest it create stigmatization and discrimination of groups of individuals. This threat, which is certainly not peculiar to genetics, has often been perceived in the past – for instance, in attitudes to disability, rare diseases, HIV infections and many other conditions. In genetics, stigmatization and discrimination could occur on the basis of the genotype or ethnic group. Population groups who are "genetically disadvantaged", for example because their genotype exposes them to the risk of developing a specific pathology or resistance to a certain medicine, could be stigmatized or discriminated against (with regard to such issues as access to health insurance, mortgages and employment). Since many points are still relatively unclear in this field and there is a lack of clear regulatory provisions, the actual risks are at present difficult to determine. The following considerations should afford some guidance.

Every individual has genes creating susceptibility to important diseases, and in the vast majority of cases genotyping does not make it possible to predict with certainty that the disease will actually occur. For instance, a person who is APOE•4-positive has an increased risk of developing Alzheimer's disease, but an APOE•4-negative subject's genotype may indicate susceptibility to stroke, cancer or even Alzheimer's as a result of alleles other than APOE•4 or of environmental factors. There are thus no genetic grounds for considering the two subjects concerned at all different with regard to their risk of developing any disease. However, this does not mean that there is no risk of stigmatization and discrimination – an employer might decide not to hire the first person on the basis of the available information, but accept the possible risks of not having any information on the second person.

Another source of stigmatization can be the ethnic group. It is well known that specific mutations in BRCA1 and 2 genes responsible for hereditary breast cancer are more frequent in Ashkenazi Jewish women than in other ethnic groups, including Sephardic Jewish women. This means that if a woman is known to be of Ashkenazi Jewish descent, she is implicitly known to have an increased risk for early development of hereditary breast cancer, even if the result of her genetic test is not known. Similar considerations apply to the whole range of environmental risk factors, for example in the case of populations living in highly polluted areas as opposed to the countryside, or of a Sephardic Jewish woman who may be susceptible to other diseases or to the sporadic form of breast cancer, far more common than the hereditary type. The EPO (European Patent Office) nevertheless granted a patent to the University of Utah Research Foundation on a BRCA2 gene mutation analyzed for "in vitro diagnosis of predisposition to breast cancer in Ashkenazi Jewish women". The European Society of Human Genetics, which is opposed to the patent, regards it as discriminatory against the Ashkenazi Jewish ethnic group. See the text-box "***Patents on hereditary breast cancer genes***", in the section "***Commercial and patent rights***".

The possibility of discrimination against patients with a genotype predisposing to a reduced drug response is sometimes raised. This information cannot harm the patient – indeed, it is surely useful to know in advance if the patient will experience benefits or adverse effects from a medicine, rather than find this out only when treatment is in progress. It is sometimes thought that systematic use of pharmacogenomics and pharmacogenetics in drug research and development could channel research toward drugs for common responder genotypes, to the detriment of medicines for less frequent genotypes. However, even without pharmacogenetics, a pharmaceutical company developing a new drug will have access to the same sort of information on the proportion of responders and non-responders once early phase II studies have been carried out.

The risks of stigmatization and discrimination on genetic grounds thus seem to be based more on overestimation of the predictive power of genetic testing and genetic predisposition, rather than on valid scientific and medical assumptions.

Types of human genetic research and associated risks

Genetic research is characterized by the collection and/or analysis of biological material from which genetic material is extracted (DNA, RNA, proteins) and used in genetic or genomic analyses for purposes of research, not of diagnosis.
Genetic studies designed and submitted to ethics committees can be classified according to their objectives and the methodology used for analysis. There being no official or internationally approved terminology, **the working group proposes the following definitions and has used them in the present document.** They reflect current usage in the scientific community, but may not correspond to that of other documents, for example EMEA's 2003 "Position Paper on Terminology in Pharmacogenetics"[(11)].

Research on disease genetics (often referred to as "genetic research") encompasses genetic studies in humans to investigate the influence of single genetic factors on the development, severity and progression of a disease, through the collection and analysis of DNA, RNA and proteins. If research with these aims studies DNA or RNA to analyze many genes, large chromosomal regions or the whole human genome, the study is referred to as *genomic research* or, if proteins are involved, *proteomic research*.
Pharmacogenetic research is defined as genetic research in humans to study the influence of single genetic factors on response to drugs, through the collection and analysis of DNA, RNA and proteins. If research with these aims studies DNA, RNA or proteins to analyze many genes, large chromosomal regions or the whole human genome, the study is defined as *pharmacogenomic research* (see **Table 10:** *Terminology in genetic research*).

Table 10: *Terminology in human genetic research*

Category of research	What is studied	What is analyzed
Genetics	*The influence of single genetic factors* on the development, severity and progression of a *disease*	DNA, RNA and protein variations in single genes
Genomics	*The influence of genetic factors* on the development, severity and progression of a *disease*	DNA, RNA and protein variations in multiple genes, large chromosomal regions, or in the whole genome.
Proteomics	*The influence of genetic factors* on the development, severity and progression of a *disease*	Protein variations
Pharmacogenetics	*The influence of single genetic factors* on *drug response*	DNA, RNA and protein variations in single genes
Pharmacogenomics	*The influence of genetic factors* on *drug response*	DNA, RNA and protein variations in multiple genes, large chromosomal regions, or in the whole genome.

Information produced by studies on multifactorial traits entails a more limited risk of generating discrimination than that from studies on Mendelian traits. Even in the unlikely event of all the genes associated with susceptibility to a specific disease being known, there would still be no certainty of the disease developing. Lifestyle can also affect the likelihood of this happening, modifying the proportion of total risk related to the environment and not to genes.

The information produced by a susceptibility test for a given disease, which involves a greater potential risk of discrimination or stigmatization, requires greater data protection than data from a pharmacogenetic test. The potential psychological implications for subjects or kin are also greater in a genetic research than pharmacogenetic research, in which subjects participate with an already established disease diagnosis and with an ongoing treatment. (See **Table 11:** *Risks of genetic and pharmacogenetic analyses*).

Table 11: *Risks of genetic and pharmacogenetic analyses* .

Genetic research	Pharmacogenetic research
• Greater risk of discrimination and stigmatization • Greater potential implications for insurance and employment • Greater need for data protection • Greater psychological impact (morbidification) on the subject and kin • Difficulty of understanding: the meaning of probability and risk is not very clear	• Not diagnosis or prediction of disease, but prediction of response to a specific drug • Subject with an already diagnosed disease or risk factor • The test is performed when a proven treatment for the disease already exists
THE PRINCIPLE OF AUTONOMY PREVAILS	**THE THERAPEUTIC PRINCIPLE PREVAILS**

Sometimes genetic study protocols have various objectives, for instance related to pharmacogenetic and genetic research. The database from a clinical trial on a drug can include, besides information on drug efficacy and tolerance, data on disease progression rate in untreated individuals or on different levels of severity (e.g. data collected in basal conditions or from the placebo group). These phenotypic data are an excellent resource both for pharmacogenetic analysis and for the study of genetic influence on disease severity or progression rate. In pharmacogenetic studies on drugs used for the treatment of diseases requiring biopsies or surgery (e.g. tumours), it is often possible to include a genomic or pharmacogenomic evaluation, given the availability of pathology samples and the scientific value of their analysis.

It should be underlined that it is not always possible to distinguish the type of information generated by a specific test. It is now well known that a gene encodes for the synthesis of an average of ten proteins, which may be involved in different pathways and pathologies. For example, the APOE gene is related to the risk of cardiovascular pathologies and Alzheimer's disease, while the ALOX5 gene has been associated with development of cardiovascular pathologies and response to asthma treatments. At present there are not many examples of a single gene playing a role in different pathologies, but in the future, with the progression of knowledge on genes and their role, the extent of this phenomenon is likely to prove far greater. If a gene can have several functions, this means that it may carry a range of different information with varying levels of ethical, legal and social risk, none of which can be isolated from the others.

The document "Recommendation of the Committee of Ministers to member states on research on biological material of human origin"[3] recommends that the risks associated with research using biological material must not be disproportional to the potential benefits of the research.

Benefits of human genetic research

Currently most genetic research consists of exploratory studies involving low penetrance genes, so that there is little likelihood that results will be of immediate clinical use for subjects. There is thus no direct benefit for the subject who participates in genetic research. There may certainly be a benefit for the community, since the research may contribute to scientific discoveries or to possible applications in the medical field. Given the lack of direct benefits, it is appropriate that if sample collection occurs within a clinical study, the subject be given the chance to participate in the clinical study (which will supposedly offer direct benefits), not the genetic study. The informed consent must clearly indicate that direct benefits are not expected.
However, the first non-exploratory pharmacogenetic studies are now in progress and may generate results potentially useful for the health of the subject. The informed consent must in this case specify whether results useful for the individual are expected and indicate how they will be disclosed to the subject.

Clinical research and genetic research in humans: similarities and differences

Clinical research is one of the best described processes, subject to specific and internationally shared regulations. On the other hand, there is a lack of regulations and even of official guidelines or international consensus on genetic research in humans. To what extent regulations and experience in the field of clinical research can be transferred to genetic research is thus a legitimate question.
The principle that has inspired the creation of norms for clinical research has been the protection of the subject participating in the study from potential physical damage due to the procedures or drugs involved. Such damage is not likely in genetic research, because the most invasive procedure involved is collection of blood samples. It should also be realized that studies of genetic influence on disease do not involve drug treatments, and even pharmacogenetic studies usually consist in the collection of DNA samples within clinical studies which do include drug treatments.
The working group is in agreement with the widely held position that all the regulations and principles valid for clinical research are also applicable to human genetic research. However, these are not enough in genetic research, because they do not address important issues typical of such research, such as storage and the use of collected samples. It is therefore necessary that genetic research be regulated by guidelines and specific norms, based on those for clinical research appropriately complemented by other specific indications.

Summing up and recommendations

The potential risk associated with human genetic research regards “information” deriving from the data collected during the study or produced by genetic analysis. This risk also exists in non-genetic studies.

The potential “information-related” risk associated with pharmacogenetic research is normally lower than that of genetic testing in multifactorial disease, which is in turn lower than that of genetic research on a monogenic disease.

Genetic research today normally benefits the scientific and medical community as well as society at large. Research contributes to scientific knowledge and its possible future clinical application. To date, genetic research has rarely resulted in direct benefits for the patient, but the first pharmacogenetic studies are in progress and will generate information potentially useful for the health of the individual who has participated in the research.

All the rules and principles valid for clinical research are applicable to genetic research, but other specific issues must also be considered – e.g. management of biological samples (DNA, RNA).

In the drafting or assessment of clinical research protocols in human genetics, the working group recommends that the following points be considered:

- **the element of risk associated with human genetic research is mainly related to information collected during the research or produced by genetic analysis**
- **the magnitude of this risk depends on the type of research and the character studied**
- **to counter this risk research must adopt proper measures for data protection and for management of individual information**
- **to date it is the community rather than the individual which benefits from human genetic research, but studies which may benefit patients are already in progress and could become more frequent in the near future**
- **risks and benefits of the specific research must be adequately described in the informed consent.**
- **every human genetic study must use the same rules and principles which regulate clinical research, and must also appropriately address typical issues like those pertaining to management of DNA samples (DNA, RNA).**

Data protection

Application of Law 196/2003 to genetic research

Data protection for the subject participating in the research is not only an ethical duty but also mandatory by law, the Italian implementation of the 1995 European Directive on data protection issues (Directive 95/46/EC) being Law 196/03.

In Italy, personal data protection in genetic research is subject to the following laws and regulations:

- Law 196/03, approved on 30th June 2003: "Code concerning the protection of personal data", commonly referred to as the "Data Protection Code" or "Data Protection Act" [(19)]
- Authorization n° 2/2005 for management of data revealing an individual's health and sex life [(20)]
- General ruling of March 31st 2004, regarding situations exempt from the duty of notification [(21)].

To understand how the regulations for the protection of personal data are applied to genetic research, it is first necessary to clarify how the different types of data are defined in the Data Protection Code. Art. 4 of Law 196/03 specifies the following definitions:

- **Personal data:** any information related to a physical person, juridical person, corporation or association who/which is identified or identifiable, even indirectly, by reference to any other information, including a personal identification number;
- **Identification data:** personal data that allow direct identification of the subject;
- **Anonymous data:** information that originally or after handling cannot be associated with any data belonging to an identified or identifiable subject. These data are exempt from the requirements of Law 196/03.

The Code states that anonymous data are exempt from the requirements of Law 196/03. But which types of data handled for research purposes does the Code consider anonymous?

It should be noted that the categories of data described in Art. 4 do not correspond directly or immediately to those identified in the EMEA document "Position paper on terminology in Pharmacogenetics"[(11)]. Data the EMEA document includes in the category "complete identification" are classed as "personal data" in Art. 4 and, being non-anonymous, are subject to the requirements of Law 196/03. Data included in the categories "anonymized" and "anonymous" in the EMEA document are referred to as "anonymous data" in Art. 4 and are therefore not subject to the requirements of Law 196/03.

Data included in the categories "simple coding" and "double coding " in the EMEA document (coded data) were controversial in the past. One interpretation was that these data could be considered anonymous according to Law 196/03, on condition that any person handling them could not trace the subject's identity other than by extraordinary, disproportionate and unreasonable means. According to this interpretation, coded data handled by a sponsor are anonymous and therefore not subject to the Data Protection Act, on condition that only the investigator, and not the sponsor, can link the code to the name of the subject. Another interpretation is that coded data could in any case not be considered anonymous if any individual other than those actually involved in handling the data were able to trace the subject's identity. According to this interpretation, coded data handled by a sponsor are not anonymous and therefore are subject to the Data Protection Act, since the investigator – though not the sponsor – can trace the subject's identity.

The document "Recommendation of the Committee of Ministers to member states on research on biological material of human origin", drafted in October 2005 by the European Council Committee of Bioethics and adopted without modification by the Committee of Ministers of the European Council on 16th March 2006 [(3)], clarifies this issue. Art. 3 of the text states that any biological material identified through a code must be considered "linkable anonymized material" if the code is under the control of a third party and the user of the biological material does not have access to it.

As a result, coded data handled by a sponsor who has no access to the subject's name (since only the investigator can link the code to the subject's name) are considered anonymous. In this case the

sponsor is not the person in charge of handling the data collected or produced by the research, and the data are exempt from the Data Protection Act.

Requirements of Law 196/2003 for genetic research

Law 196/03 states that genetic research involving non-anonymous data is subject to the requirements described and summarized in **Table 12:** *Requirements of Law 196/03 in relation to the nature of data handled.* These requirements are set out in the following subsections.

Authorization of the National Data Protection Office and general authorizations

For the handling of sensitive data, the Code specifies the need to apply for the authorization of the National Data Protection Office before starting specific data processing.
To regulate various types of data processing uniformly, the National Data Protection Office has issued six general authorizations for different types of sensitive data.
It is not necessary to submit an application to the National Data Protection Office if the data management is covered by one of the six general authorizations – in such cases it is enough to identify the type of data concerned and follow the relevant indications.
The investigator who handles the subject's personal data for research aims (non-anonymous data) does not have to apply for the National Data Protection Office's authorization to handle genetic information collected or produced during the study, since general authorization n° 2/2005 (1.2 *(a)*) applies in such cases.

Notification to the National Data Protection Office

According to Art. 37, the National Data Protection Office must be informed of the handling of data involving genetic information.
The National Data Protection Office's general ruling of 31st March 2004 (regarding situations exempt from the mandatory notification requirement) has clarified a number of issues regarding mandatory notification to the National Data Protection Office, the following types of data handling being considered exempt from this obligation:

> *non-systematic handling of genetic data [....] by healthcare professionals, also in cooperation with others involved, regarding data not accessible by a third party through computer terminals. The scope of this ruling is limited to data and operations, including communication, essential to the objectives of health protection or the physical safety of the subject or of third parties.*

If the category of data handling falls within this definition, it will not be mandatory to notify the Data Protection Office. In all other cases, notification as required by Art. 37 of the Code is mandatory.
In the case of genetic research, the above document does not exempt investigators from the obligation to notify the Data Protection Office of data handling, since genetic data handled for research purposes do not fall within the exempt categories listed.

Obtaining the subject's informed consent

According to Art.13, informed consent is required to ensure the subject knows that the person in charge of data handling:
• is collecting data;
• will handle such data only for the aims expressly stated in the informed consent;
• will store the data in a databank;

• will be able to update and correct data if the subject exercises the Right of Access specified in the Code and objects to data handling or asks which data have been acquired by the person in charge of handling.

By giving informed consent, the subject confirms to the person in charge of data handling that s/he wishes personal data to be handled for the aims stated in the informed consent sheet.

If the subjects asks for withdrawal of the consent or cancellation of personal data, the person in charge of data handling can keep the collected data only if they have been anonymized (preventing any possible future link between the code and data, even if owned by third parties).

The sponsor managing "anonymous" data as defined above is not obliged to cancel the data, even in the event of the subject's consent being withdrawn. However, the sponsor must not generate additional data after withdrawal of the subject's consent. In practice, at least in research by industrial sponsors, data produced before withdrawal of consent are never cancelled from the database because they are of legal and scientific importance for documentation of the research. Indeed, such is the value of human study databases that after their official completion they are protected by systems to prevent cancellation of data by any user.

Table 12: *Requirements of Law 196/03 in relation to the nature of the data handled.*

Nature of data as defined by Law *196/03*	**Non-anonymous**	**Anonymous**
Category	Identifiable, directly or through a code (coded) if whoever uses the biological material can link the code to the subject's identity.	Coded if whoever uses the biological material cannot link the code to the subject's identity because the code is managed by a third party (anonymized linkable). Not identifiable (anonymized non-linkable) if it is not possible by reasonable means to trace the subject's identity.
Requirements for the study promoter	The promoter of the research is in charge of data handling and thus subject to all the requirements of the Data Protection Act.	The promoter of the research is not the person in charge of data handling and not subject to the requirements of the Data Protection Act. In general terms, data protection is still an obligation.
Requirements for the investigator	S/he does not have to apply for the authorization of the National Data Protection Office, since a general dispensation (authorization 2/2005, 1.2 lett. *a*) applies in such cases. The National Data Protection Office must be informed about the handling of genetic data. The subject must be informed about the collection of personal data, the aims of data handling, and the right to withdraw and have data cancelled. Informed consent must be obtained. Data must be stored for the period required by law or regulations, but in any case not longer than strictly necessary for the aims of the research.	S/he is not subject to the requirements of the Data Protection Act. In general terms, data protection is still an obligation.

The ethical duty to protect data

In addition to legal obligations, there is also an ethical duty and certain measures can be adopted for the protection of a subject's data, including:

- labelling of test-tubes and data collection sheets only with codes or other indications that make it impossible to identify the subject;
- choosing the appropriate level of identification of samples and data (as detailed in the paragraph "***Level of identification***");
- ensuring that the subject's identity is known only to the investigator and co-workers involved in the study;
- ensuring that access to archives (paper or electronic) and to rooms where the biological samples are stored is limited to staff involved in the research;
- for exploratory genetic research without the expectation that results will be clinically useful for the individual, ensuring that participation in the study and any data produced by the genetic research are not recorded on the clinical record form; information collected during the research, including informed consent, should be collected separately in a safe archive distinct from the clinical record;
- providing results of genetic analysis only to patients who ask for them and to no one else (unless the patient gives specific consent), or only to the investigator in the case of results which are clinically useful for the patient;
- describing in the protocol data protection procedures and methods for pseudo-anonymization (coding) or anonymization (temporary or irreversible) of data;
- entrusting sample collection, management of biobanks and all procedures involved only to qualified and competent staff;
- transferring data and samples among different centres only when necessary or appropriate for the aims of the research, and subject to individual consent.

Summing up and recommendations

Data protection for the subject participating in the research is both mandatory by law (Law 196/03) and a more general ethical duty. Non-anonymous data handled for a genetic research are subject to Law 196/03, as stated in Art. 4. These data include identifiable data or coded data when whoever uses the biological material (e.g. the sponsor) can link the code to the identity of the subject. Data are anonymous and thus not subject to the Dat Protection Act when they are anonymized and not linkable to the identity of the individual, or when they are anonymized and linkable but whoever uses the biological material cannot link the code to the subject's identity. For more detailed information about this topic, see ***"Level of identification".***

The investigator who handles the subject's personal data for research aims (non-anonymous data) does not have to apply for the National Data Protection Office's authorization to handle genetic information collected or produced during the study, since general authorization n° 2/2005 (1.2 *(a)*) appliesin such cases. However, the National Data Protection Office must be informed of the handling of data involving genetic information.

In the drafting or assessment of clinical research protocols in genetics, the working group recommends that the following points be considered:

- **mandatory compliance with Law 196/03 according to the type of data handled**
- **in addition to the requirements of Law 196/03, there is a more general ethical duty of data protection**
- **some additional measures can prove valuable for purposes of data protection – e.g., ensuring that access to archives (paper or electronic) and to rooms where the biological samples are stored is limited to staff involved in the research, limiting access to genetic**

analysis data so that they are available only to patients and, in the case of results which are useful for the patient's health, to the investigator.

Study rationale and aims

As a general criterion, as for any other biomedical research, every proposed genetic study must be supported by a rationale and its aims must be clearly described. Sample collection without a clear purpose, or for unspecified future uses should not be authorized. Whenever possible, the protocol should refer to similar studies already completed. However, genetic research is a relatively recent field and proposed research could present a totally new design and rationale. In such cases it is important to balance the lack of information on previous studies with the consideration of what is new in the research. In this way it can be ensured that the most innovative studies will not be penalized, while at the same time research lacking a serious scientific basis or not providing subjects with sufficient information for appropriately informed consent will not be possible.

Limiting the freedom of genetic research

Sometimes when writing a research protocol it is not possible to describe thoroughly and precisely the study rationale, and the aims can be expressed only in general terms. This is more and more often the case, and is likely to be even more common in future. This is because industrial pharmacogenetic research is becoming increasingly concerned with the influence of genetics on the efficacy and tolerability of new drugs in development, to improve knowledge of efficacy and tolerability so that decisions related to drug development can be properly supported.
For this purpose, samples of DNA (or other material) are collected during the early phases of a new drug's development, when knowledge of it is limited. Sometimes the study rationale is generated after the conclusion of this research. For example, if DNA samples are collected during the initial phase II study performed on a new drug in development, at the time of writing the protocol it is not predictable whether there will be individual differences in efficacy and tolerability as a basis for a pharmacogenetic evaluation. At the end of the study, some variations in efficacy might have been observed, but with only very limited numbers of adverse events and little information on them. A subsequent study on the same drug may identify individual variations in relation to adverse events, providing a rationale and detailed indications for analysis of previously collected samples.
See "**Storage**" and "**Use of samples**" for information on the storage and use of samples after the conclusion of the study.

The opinion of the working group is that the collection of biological samples without a rationale and a clear purpose, or for unspecified future uses, should not be authorized. The study rationale and aims must always be stated.
The working group thinks that the most innovative research should not be discouraged. When a specific study's rationale and aims cannot be described in considerable detail when the protocol is written, it is enough to state:

- **the general rationale for sample collection**
- **why it is not possible to state the rationale of the study concerned**
- **the aims of the research, at least in general terms**
- **cases in which samples will or will not be analyzed**
- **the guarantee that, in the event of initial prospects changing, subjects will be contacted for explicit consent.**

It is important to consider how much detail is required in defining the aims of the research when they cannot be stated exhaustively at the time of writing the protocol. The following examples can be taken as an illustration in this regard:
• *to study the effect of genetic variants on pharmacokinetic and pharmacodynamic parameters and on the tolerability of drug X*
• *to study the effect of genetic variants on susceptibility, development, progression and severity of disease Y.*

Summing up and recommendations

As a general criterion, as for any other biomedical research, every proposed genetic study must be supported by a rationale and its aims must be clearly described.

Sometimes when writing a research protocol it is not possible to describe thoroughly and precisely the study rationale, and the aims can be expressed only in general terms. This is more and more often the case with the most recent research.

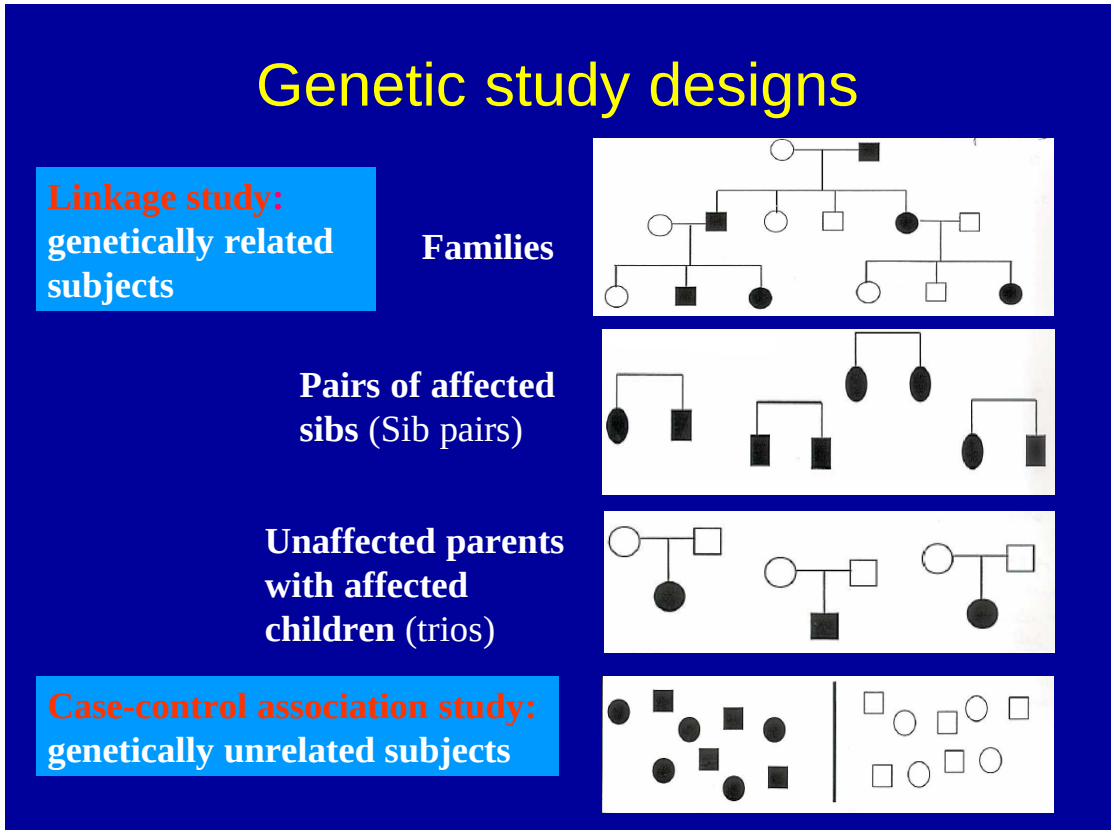
In the drafting or assessment of clinical research protocols in genetics, the working group recommends that the following points be considered:

- **the study rationale and aims have always to be described**
- **if it is not possible to describe the study rationale and aims in considerable detail, they must be described at least in general terms.**

Study design

Irrespective how a clinical study which may include genetic research is designed (cohort study, transverse study, population study, case-control study, etc.), two general methods are used for genetic research: case-control association studies and linkage studies (family segregation studies). In case-control association studies, genetically unrelated individuals are (usually) recruited, for instance diabetics (cases) and non-diabetics (controls), responders (cases) and non-responders (controls) to a certain drug.
In linkage studies, genetically related individuals are recruited: families, pairs of affected siblings (sib-pairs), unaffected parents and affected children (trios). See **Figure 2:** *Genetic study designs.*
Linkage studies in families are mostly used to study Mendelian diseases and are useful for initial identification of a genomic region correlated to a phenotype, but are limited by the need to have many, if possible very large families.
Case-control association studies are typically used in pharmacogenetic and genetic epidemiological studies. This approach will become more and more common as a result of recent advances in SNP haplotype analysis on chips[22]. In association studies it is necessary to take into account possible confounds, such as stratification as a result of differences in ethnic group, age, sex, and other phenotype-related factors.

Figure 2: *Human genetics study designs*



Summing up and recommendations
Study designs are of two main types: *case-control association studies*, recruiting genetically unrelated individuals, and linkage studies, involving genetically related individuals (families, sib-pairs).

Linkage studies in families are mostly used to study Mendelian diseases and are useful for initial identification of a genomic region correlated to a phenotype, but are limited by the need to have many, if possible very large families. Case-control association studies are typically used in pharmacogenetic and genetic epidemiological studies, and it is necessary to take into account possible confounds, such as stratification as a result of differences in ethnic group, age, sex, and other phenotype-related factors.
As a general criterion, as for any other biomedical research, the design of the proposed genetic study must be appropriate to achieve the aims of the research.

In the drafting or assessment of clinical research protocols in genetics, the working group recommends that the following points be considered:

- **the design of the study must be appropriate to achieve the aims of the research**
- **linkage studies in families are mostly used to study Mendelian diseases and are useful for initial identification of a genomic region correlated to a phenotype, but are limited by the need to have many, if possible very large families**
- **case-control association studies are typically used in pharmacogenetic and genetic epidemiological studies, and it is necessary to take into account possible confounds, such as stratification as a result of differences in ethnic group, age, sex, and other phenotype-related factors.**

Sample size

In principle, every genetic study, like any other biomedical research, must involve a sample of suitable size to achieve the stated aims with the desired power. In practice, sample size in genetic research often depends on factors that are unknown at the beginning of the study, such as the penetrance of the gene and the frequency of the polymorphism in the studied population. Since attitudes to this aspect of modern genetic research are subject to rapid change, no definitive indications can be offered.

For a more detailed analysis of this issue, the working group suggests some valuable bibliographic references regarding the different approaches to research [23] and genomic association studies[24, 25].

The literature reports examples of effects shown even with no more than 30 cases and 30 controls, but it should be borne in mind that the possible "publication bias" of small studies with statistically significant results – as compared to studies with negative results – could give an unreally optimistic picture of sample size requirements. A study with a small sample is often accepted for publication if it has succeeded in showing an effect, while it is rejected for insufficient sample size if no correlation has been identified. It is generally true that genetic research with a small sample size (e.g. less than 100 subjects) has little likelihood of showing any effect. If 100 subjects are recruited for a genetic study, they will be divided into 3 groups according to genotype (e.g. AA, AG, GG) and the size of each group will depend on the frequency of that genotype. If the studied genotype, for example GG, has a low frequency (for instance 5%), the GG group will include 5 subjects and it will be impossible to draw any conclusion from this study. On the contrary, given a studied genotype AA with a 70% frequency, it will be possible to draw conclusions from the research if the effect of the polymorphism on the studied feature is important. The situation is even more complex because a study often analyzes up to several dozen genes, each with a number of polymorphisms which can determine considerable variation in the frequency and importance of their effect. It is not always possible to establish sample size with a view to the aims of the genetic analysis. For example, in a pharmacogenetic study within phase II research on a new drug, sample size usually corresponds to the aims of the clinical part of the study; it is thus necessary to ensure that it is also appropriate for pharmacogenetic evaluation. In such situations, if the research has a solid scientific rationale, it is possible to select genes for analysis by a variety of means. One possibility is to use the frequency of polymorphisms as a selection basis, while another is to bear in mind that research can demonstrate only important genetic phenomena and run combined analyses of samples or data from different studies with compatible phenotypes.

Although the need to avoid studies with a clearly inadequate sample size is recognized, the working group believes that a study with a solid rationale should go ahead and that any limitations related to sample size should be addressed by the above-mentioned methods. If accurate assessment of the necessary sample size is not possible, the grounds on which the suggested sample size is considered appropriate must be stated.

Summing up and recommendations

In principle, every genetic study, like any other biomedical research, must involve a sample of suitable size to achieve its stated aims.

In practice, sample size in genetic research often depends on factors that are unknown at the beginning of the study.

In the drafting or assessment of clinical research protocols in genetics, the working group recommends that the following points be considered:

- **the study protocol must contain a justification of the proposed sample size**

- **if accurate assessment of the necessary sample size is not possible, the grounds on which the suggested sample size is considered appropriate must be stated.**

Biological samples: level of identification, storage and use

Biobanks and collections of biological samples

To correctly outline the regulatory framework to which samples collected during genetic research are subject, it is useful to distinguish human tissue banks (biobanks) from collections of human biological samples generated within genetic research for the sole aims of the research concerned.
The term biobank has been defined in various ways and there is still no general consensus on the correct definition. Council of Europe Recommendation R (94) 1, dated 14th March 1994, defines the human tissue bank as a non-profit organization that must be officially recognized by the health authorities of member states, and must guarantee the treatment, storage and distribution of the material. This distinction has important practical implications, such as the authentication and the accreditation to which biobanks, but not biological material collections, are subject.
The simple collection of biological samples within a genetic research setting is similar to the collection of biological materials referred to in the above-mentioned document and differs from a biobank in that the latter has broader aims. In particular, one of the main purposes of the biobank is to make biological samples available to those who apply for permission to analyze them.
Some principles, such as the need to guarantee correct storage of samples, apply both to samples from a biobank and those from a collection of biological samples made during a specific study. Other more specific issues need to be managed differently. As an example, unlimited storage of biobank samples is in some cases justifiable, while this is difficult for a collection of samples to be used in a study. The level of detail required regarding use of samples can also differ in the two cases.
In the present document, the term "biobanks" is used as suggested in the above-mentioned "Recommendation" and "Guidelines" documents, while the expression "collection of biological material for research purposes" is used to indicate material collected within a genetic study for the sole purpose of carrying out the study concerned.
The documents "Genetic Biobanks – Guidelines", issued by the Italian Society of Human Genetics in conjunction with Telethon, and "Guidelines for the institution and the accreditation of biobanks", by the National Committee for Biosafety and Biotechnologies, supply detailed recommendations on the aims, setting up, management and the accreditation of biobanks [(26), (4)].
The European Society of Human Genetics has issued some recommendations on collections of biological samples, presented in the document "Data Storage and DNA banking for biomedical research: technical, social and ethical issues" [(27)].
Biological samples collected during research, if stored, constitute a collection of biological samples. See the text-box ***"Collections of biological samples and genetic and clinical databases".***

Collections of biological samples and genetic and clinical databases

The purpose of any genetic study is to find an association between a genotype and a phenotype. For this reason it is necessary to have full availability of clinical data regarding every collected sample and to associate them with those obtained from genetic analysis of the collected samples. Biological samples collected for research purposes thus have the peculiarity of making up a collection of both samples and data. Such data are collected in clinical phenotype databases and genetic databases. The clinical phenotype database consists of a collection of all the available phenotype information, containing demographic and medical information – e.g. age, presence or absence of a certain disease, results of clinical examinations on parameters such as glycaemia.
A genetic database is the total information originating from the analysis of the collected samples. The data contained in the genetic database must be related to each subject's phenotype data. Usually the subject is identified in databases by a numbered code and not by name, so that those managing and operating on the database do not have access to the subject's name.

Some large population studies which have recently started have given rise to collections of biological samples and in some cases to biobanks which are larger and serve broader purposes than normal collections of biological samples created for a more specific study. These collections include samples and information regarding whole populations. They are linked to genotyping projects started in some countries and sometimes very different criteria must be applied to their management, with many different levels of protection for samples, data and donors. See the text-box ***"Some large population studies"***.

Some large population studies

The Human Genome Project, the HapMap project for haplotype mapping and other studies have increased the need for genetic data on vast, well characterized populations. Such studies can contribute greatly to understanding how genetic and environmental factors influence the development of common diseases.

These broad genetic epidemiological studies require huge scientific and organizational resources, calling for close collaboration between different research groups and projects. To facilitate this and coordinate efforts between different research groups and projects involved in population genomics studies, the "Public Population Project in Genomics (P^3G)" was set up in Canada. This is an international non-profit organization to which the organisers of most ongoing population genetics projects in progress belong [28]. The ain of P^3G is to bring together the efforts of different research groups and pool study databases, creating a sort of worldwide population study project.

The following paragraphs present highlights of some large-scale ongoing population studies.

Iceland: **Health Sector Database** **(http://www.decode.com;http://www.mannvernd.is)**

- A project of deCode Genetics, approved in 1998 by the Icelandic Parliament with the consent of about 70% of the population
- Objectives: to acquire new knowledge on diseases and health, to improve the quality and economy of the health system, to develop a high-technology industry in Iceland favouring the employment of highly specialized resources, to attract investments
- deCODE Genetics, which is a private firm, was granted a licence in 2000 to file genetic and medical data of 275,000 Icelanders over a 12 year-old period
- Consent: presumed, the will not to participate needing to be expressed within six months, meaning that deceased individuals and minors had no choice in the matter. The faculty to withdraw was not allowed once data were anonymized
- Data protection: no protection vis-à-vis commercial and government users
- Exclusive rights of use: monopoly of deCODE, difficult for independent researchers to compete.

Quebec, Canada: **The CART@GENE Project** **(http://www.rmga.qc.ca)**

- A project by the researchers of the Réseau de Médecine Génétique Appliquée (RMGA).
- The project plans to create the first genetic map of Québec with the following objectives:
  - to identify the genetic origin of complex diseases such as cardiovascular and psychiatric diseases etc.
  - to identify protective genes against current diseases
  - to direct research toward new treatments
  - to make better use of therapeutic and healthcare resources within the community
  - to help historians and sociologists understand population movements and migrations
- The project relies on the voluntary participation of 50,000 Quebec inhabitants, aged 25 - 74, making up a representative sample of about 1% of the entire population.

United Kingdom: **UK Biobank (http://www.ukbiobank.ac.uk)**

- A project by Medical Research and the Wellcome Trust
- Purpose: to test the DNA of about 500,000 volunteers aged 45 - 69, distributed according to a regional criterion, to understand:
  - the influence of particular genes on the course or severity of a disease
  - the number of patients with a specific genetic mutation, who may have greater predisposition to a specific disease.
  - the influence of interaction between genes and environmental factors on the causes of a disease
- Data protection: separation between registry and genetic information
- Consent: written informed consent specifying the different research issues; withdrawal possible at any time
- Access to data: possible for a subject to have personal health information. Private insurance companies and other organizations not to have access to information
- Dissemination of global results on Internet.

Estonia: **Estonian Genome Project (http://www.geenivaramu.ee)**

- A June 2000 project developed by the Estonian Genome Foundation and authorized by Parliament through the Human Gene Act
- Collects data and samples of about 100,000 participants to set up a database including information on health status, DNA, plasma and genealogical data
- Purpose: bank of phenotypic and genotypic information to allow research on the health status and genes of about 1 million Estonians (70% of the entire population), storing tissue samples to isolate the genes causing or affecting the most common diseases
- Exclusive rights of use: the Estonian Genome Foundation, responsible for the Project, is the owner of the database
- Information processing rights are delegated by the Foundation to the private company EGeen Inc, which supports the project in compliance with legal requirements
- Consent: any individual is free to decide whether to participate in the Project. It is possible to withdraw at any time, waiving access to personal genetic information in that withdrawal entails definitive cancellation of the identification of the samples concerned
- Access to data: the genetic data bank can fulfil both scientific and personal purposes: if the donor wants specific information, s/he will receive it free of charge
- Confidentiality: ensured by separating personal and genetic information; they can be matched by a specific code, known only within the Foundation and usable at the request of the donor or personal physician.

United Kingdom: **Centre for Integrated Genomic Medical Research (CIGMR)**

Includes families and various populations with a specific disease, producing data over a period of time for more than 20,000 individuals.

Central Europe: **Danubian Biobank Foundation**

Collects phenotypes, genotypes and biological samples of cohorts and groups at risk in 6 countries of Central Europe. Clinical endpoints defined to help identify and validate targets and biomarkers for the common problems of aging.

European Community**: GenomEUtwin**

Studies 600,000 pairs of twins within a partnership of European registries, matching genetic, epidemiological and phenotype data related to common diseases.

Germany: **KORA-gen**

Collects phenotypes, genotypes and parameters in a bank which is constantly updated with biological samples from 18,000 subjects aged 25 - 74 years, for an epidemiological genetic study.

***Sweden*: LifeGene**
LifeGene is a prospective cohort study that will combine biological information with information on the lifestyle of 500,000 people living in Sweden, to improve the understanding of interactions among hereditary factors, lifestyle and environment in relation to the development of most common diseases.

***United States:* National Heart, Lung and Blood Institute**
The National Heart, Lung and Blood Institute supports research carried out through epidemiological studies to determine temporal and demographic prevalence, incidence, morbidity, mortality, risk factors, genetic and environmental influence and their interaction, as well as through long-term observational studies.

***Australia*: Western Australian Genetic Health Project**
The Western Australian Genetic Health Project is based on the availability of complete data about the population's health status collected in the past 30 years, including genealogical information, and focuses on collecting health data and biological samples regarding 2 million people in Western Australia.

***Japan:* The BioBank Japan Project**
The biobank contains genetic information on almost 300,000 individuals and will be useful for pharmacogenetic studies and susceptibility analyses for complex diseases and tumours.

***Gambia*: The Gambia BioBank Project**
A DNA biobank envisaging inclusion of 40,000 individuals. The first biobank in Africa.

Management of biological sample collections regarding whole populations obviously entails far more complex and critical donor protection issues than is the case with collections for specific and defined studies. These have much more limited purposes and contain less informative databases. The specificity of population biobanks is also underlined in the document "Recommendation of the Committee of Ministers to member states on research on biological material of human origin", which provides a specific definition for population biobanks[(3)].
In Italy, the issue of genetic population census studies has recently been discussed in the "Document of the National Committee for Biosafety and Biotechnologies Working Group on genetic population census studies " [(6)]. The aim of this document is to promote debate and focus attention on the need to address this issue because of the prospects it offers for the study and application of life sciences and genetic medicine.
The topic of biobank management has been discussed in various documents, including "Genetic biobanks – guidelines"[(26)] and "Recommendation of the Committee of Ministers to member states on research on biological material of human origin"[(3)]. With regard to biobanks of samples collected during genetic research, the EMEA has announced that the forthcoming guidelines entitled "Concept paper on biobanks: pharmacogenetics and pharmacogenomics"[(7)] will deal with sample collection, storage, future uses, implications concerning the level of identification, and regulatory requirements concerning samples collected during research.
These documents will certainly supply a major contribution to clarifying the important issue of the storage of biological samples collected during research.
Some considerations about the most important issues related to biological samples can be found below.

Responsibility

The promoter of the research has the responsibility to guarantee that storage, analysis and transfer of samples occur according to methods that ensure their integrity and allow their utilization for the purpose they have been collected for. All facilities that have a role in the handling and analysis of samples have to be qualified in compliance with Quality Assurance procedures. If the study promoter outsources some of the scheduled activities on samples, this implies responsibility for ensuring that the supplier acts and works in compliance with a quality system (e.g. ISO, GLP, GMP etc). This principle also applies to collaboration involving different research groups. (For further information, see ***"Use of samples"***).

Level of identification

The level of identification of biological samples and data collected or produced by the study is a measure of how directly the name of the donor can be linked to them. The levels of identification indicated below are (i) those described in the EMEA document, also adopted in the CPMP "Position Paper on Terminology in Pharmacogenetics" [(11)]; (ii) those adopted by the most recent document of the European Council Steering Committee for Bioethics (CDBI), "Recommendation of the Committee of Ministers to member states on research on biological material of human origin" [(3)].

The document of the EMEA identifies five levels, summarized in **Table 13:** *"Summary table of levels of identification of samples and data according to the EMEA"*. The levels are as follows:

1. *Identified*. The samples and data are identified with the subject's name or National Health System number. This level is applied to most routinely generated health data. In research the use of the subject's name offers no advantages over a code, while it has some disadvantages for data protection and is therefore not generally used. A very convincing case would have to be stated to justify using this level of identification in a genetic protocol.

2. *Single coded*. This is the level of identification typically used in most research on human beings. The name of the subject does not appear on the DNA sample or on genetic data, a genetic code being used. The correspondence between the code and the subject's name is known only to the investigator and staff involved in the research.

3. *Double-coded*. This is an extension of the previous level, and provides a further degree of protection. A code, used by the researcher during the study, identifies the subject and clinical data, while samples and data produced by analysis are marked with a different code, known only to those instructed to perform the analysis but not to the investigator. These two codes must be matched to trace the subject's identity. The key to matching these two codes is held by a third party. This system increases the degree of protection, and the responsibility to keep the subject's identity secret does not rest solely on the experimenter. It could nevertheless be difficult to manage this system in practice, especially in cases where the collected samples are stored for long periods of time. In such cases it could be difficult to find a third party, other than the investigator and study promoter, able to guarantee safe and correct storage of the code for long periods even in the event of organizational, staff or other changes. This system is therefore little used in the practice when the samples are stored beyond the end of the study.

4. *Anonymized*. This means that samples and data are previously coded with a simple or double code, for which the connection with the subject's name has been permanently destroyed. For instance, samples and data can be collected in coded form and anonymized before analysis or long-term storage.

5. *Anonymous*. Anonymous samples and data cannot be limked to a subject in any way. The proper use of the term refers to samples and data collected anonymously from the outset so that there has never been a connection between the samples and the subject, and to samples associated only with general population information (for instance, the indication that the sample has been collected from a diabetic subject) but with no specific demographic or clinical information (for instance age, sex, cholesterol levels) that could in some way make it possible to identify the subject. In most genetic studies this level of identification is not used, because it is more important to know as much as possible about the subject in order to obtain accurate characterization of the phenotype.

Table 13: *Summary table of levels of identification of samp les and data according to the EMEA*

Category	**Link between subject identity and genetic data**	**Can the subject be identified for clinical controls?**	**Actions possible if subject's consent withdrawn**	**Subject informed of individual results**	**Level of data protection**
Identified (Complete)	Yes, directly	Yes	The sample can be withdrawn with immediate effect from any future use	Possible	Similar to general healthcare confidentiality
Single coded	Indirectly, via code key	Yes, as specified in the protocol	The sample can be withdrawn with immediate effect from any future use	Possible	Standard for clinical research, in conformity with principles of GCP
Double-coded	Very indirectly, through a set of two code keys	Yes, via protocol-specified procedures	The sample can be withdrawn with immediate effect from any future use	Possible	Double code offers greater data protection than single code
Anonymized	No. Key(s) identifying the link between genetic data and identity of subject destroyed	No	Sample and data not identifiable. Sample cannot be withdrawn once key is deleted	Not possible	Genetic data not linked to individuals
Anonymous	No	No	None	Not possible	Complete

The CDBI document classifies the possibility of identification of biological material, as described in **Table 14:** *Table indicating identifiability of biological material according to the CDBI.*
The document does not indicate preferred levels of identifiability, but recommends keeping biological material anonymized at the appropriate level for the specific research activities concerned, and justifying the level of identification used.

With regard to choice of level of identification, the working group believes that this choice must make specific reference to the protocol to which it will be applied, taking into account the observations set out below.

Table 14: *Table indicating identifiability of biological material according to the CDBI*

Category	Identifiable	Not identifiable
Definition	Biological material that, alone or in association with related data, permits identification of the subject concerned. This can occur directly or through a code.	Biological material that, alone or in association with related data, DOES NOT permit identification of the subject concerned by reasonable means.
Types	**Coded material** The user of the biological material can have access to the code (s/he can link it to the subject). **Anonymized linkable material** The user of the biological material cannot have access to the code, which is managed by a third party.	**Anonymized non-linkable material**

It is also important to bear in mind that regulatory authorities require data included in the registration dossier to be subject to audit, tracing back data in the dossier to the patient who generated them, to ensure accuracy and reliability. Research to be used for registration purposes must thus use coded samples and data.

- Complete identification should be avoided, unless there are very solid grounds for using it.
- Anonymity is generally not compatible with accurate characterization of the phenotype and anonymous data are useful only for epidemiological studies.
- Simple coding is often a good approach, guaranteeing adequate data protection, subject to appropriate safety measures and correct informed consent – see ***"The ethical duty to protect data"*** and ***"Individual results"***. Simple coding allows the patient to withdraw consent, ask for samples to be destroyed and to receive information about results produced by individual sample analysis.
- The same considerations apply to double coding, but it is in practice more difficult for the reasons already discussed (see ***"Level of identification"***). It should be used when there is a greater risk of breaching the confidentiality of the stored information. In such cases, it is preferable to simple coding.
- Anonymization ensures a good level of data protection, but it does not allow the subject to withdraw consent, ask for destruction of samples or receive information about the results of the analysis, and it is not applicable to data to be used for registration purposes. The European Society of Human Genetics document "Data storage and DNA banking for biomedical research: technical, social and ethical issues"[27] states that the decision to make the samples and data irreversibly anonymous should be carefully assessed and the recent EMEA document "Concept paper on biobanks: pharmacogenetics and pharmacogenomics"[3] highlights the implications of removing identification tags from

samples and data. For instance, in the event of a serious adverse reaction leading to death, there is no possibility to obtain another sample, while there is an important rationale for investigating the genetic background to the adverse reaction. This approach is thus appropriate only for basic research, where it is not possible to know in advance whether the results can be of immediate clinical interest for the subject and there is at the same time a specific need to keep information and results confidential.

Storage

How long can the collected biological samples be stored? It seems there is considerable variability in protocols submitted to ethics committees about storage times. Some protocols provide for the destruction of samples a few months after their collection, once the research is complete, while others provide for the storage of samples even for long periods of time (15 or 20 years) after the research has ended.

This variability may reflect different research needs. Some studies, which aim at extremely precise and detailed analysis of given genes or polymorphisms, offer no further potential contribution to scientific knowledge once the research is complete. In this case it is correct and justifiable that samples are destroyed at the end of the research. Other studies have more extensive aims and there is the possibility that samples could be used for further analyses even many years after their collection. In this case it is correct and justifiable that the samples are stored even for many years after they have been collected. In such cases the "end of the study" concept as such is vague, as for instance in pharmacogenetic research on new drugs in development.

If the research concerns a very rare disease it could be justifiable not to set a time limit to conservation of samples, given the difficulty of collecting them.

The working group believes that it is not suitable to state a precise time limit for sample storage, but that a criterion should be followed. This criterion is the likelihood that the stored samples could produce scientifically useful information. Destroying samples that could give additional information if re-analyzed on other occasions is neither scientifically nor ethically correct, just as it is neither useful nor ethically correct to store samples that have totally exhausted their potential to produce further results. The group recommends caution in the destruction of samples. Genetic research makes very rapid advances and it is difficult to predict what kind of new genes or new techniques could be available in the next few years. Destruction of samples denies any possibility of obtaining further scientific benefit from them and can make it pointless to have donated them in the first place. In most studies the genetic research component provides no direct benefit for the subject and sample donation must thus be seen as a contribution by the subject to research. For this reason it is essential that the potential of the samples to produce useful scientific results should be completely fulfilled.

Some examples of opportunities that can emerge from further analysis of samples even long after their collection are given in the text-box ***"Some opportunities related to (re-)analysis of samples stored for a long time".***

Some opportunities related to (re-)analysis of samples stored for a long time

- New scientific dicoveries, for instance newly identified genes or polymorphisms that were not known when the first analysis was performed
- New technological developments, for instance the possibility of running the analysis by a more powerful technology which was not available when the first analysis was performed
- A new rationale for the analysis which emerges later, for instance a side effect for a certain drug that was not previously known
- The possibility to examine newly emerged scientific evidence, perhaps brought to light by another research group
- The possibility to accumulate samples involving the same phenotype, collected in different studies and at different times.

The working group believes that, for the subject's protection, it is necessary to verify storage techniques and all information given about possible uses of samples, rather than storage time.
If storage techniques and possible uses are correctly defined, the time of sample storage should not be a risk for the subject. No matter how long the storage time, samples have to be kept in a safe place with controlled access. During storage, samples can be sent to other laboratories/researchers, but they have to be coded, anonymous, or anonymized. This might be done to allow particular investigations, provided that the subject is safeguarded to the same level (e.g. data protection) and that the same sample uses and conditions as specified in the protocol and informed consent are maintained. The document "Recommendation of the Committee of Ministers to member states on research on biological material of human origin"[3] states that samples and data should be transferred from the country in which they have been collected to a different one only if the latter assures a suitable level of protection.
The subject must have the right to ask at any time and without restriction for the destruction of the sample supplied, if not anonymized. This right should be limited to the person who actually gave consent. The sample must be destroyed soon after the request, this being documented and filed in the study archives.

Use of samples

The study protocol and informed consent must unquestionably clarify the uses to which collected samples will be put. The document "Recommendation of the Committee of Ministers to member states on research on biological material of human origin"[3] recommends that the informed consent should be as detailed as possible about the uses envisaged for samples and the options open to the patient. A direct outcome of the considerations in this document concerning the advisability of storing samples even after the study is over (see ***"Storage"***) is that samples can also be used after its completion.
To use the collected samples for purposes not included in the protocol and informed consent, it is necessary to approve a new protocol and mandatory informed consent explaining these new uses. The above-mentioned document indicates that, if the sample is anonymized and not linkable to the subject, it can be used for other purposes without the subject's consent, provided that this is not a violation of restrictions on its use previously specified by the subject. In other cases the ethics committee can allow use of samples without renewed consent only if:

- it was not possible to find the subject by all reasonable means
- the research has a critical scientific purpose that cannot reasonably be achieved using biological material for which specific consent has been given
- restrictions on the use of the sample previously specified by the subject are not violated.

It must be underlined that obtaining renewed consent can be extremely complex and possible results might be limited, especially in the case of a multi-centre international study which ended some years before. Reports suggest that in these cases far fewer subjects than originally involved give their consent for new uses of samples. The reasons are mostly organizational – an investigator changing hospital, subjects no longer referring to a particular centre because in the meantime they have moved, died or are unavailable for a variety of reasons. Sample size based on renewed consent might thus be too small or even introduce a bias into the analysis – for instance, subjects who have died might well be the ones in whom the studied disease was most severe or the drugs used were least effective.
A single-centre study or one with a limited and very specific purpose, run for just a short time, is different in this respect. Renewed consent can in all likelihood be more readily obtained in such cases.

The working group thinks that:

- **samples and data collected within a certain study can also be put to future use, provided that this is consistent with the aims of the initial study, without the need for renewed authorization or informed consent**
- **samples and data collected within a certain study can also be put to future use, even if different from the aims of the initial study, provided that the rationale and purpose of such future use are specified in the original protocol and informed consent and that the donor has given prior consent to such future use**
- **if during the storage of samples an opportunity emerges to run analyses not directly connected with the aims of the initial study, and for which the donor has not previously given consent, samples and data can be used only after obtaining specific renewed authorization from the ethics committee and informed consent (where anonymized non-linkable samples and data are not involved).**

It is not possible to write in advance a standard list of which specific (legally permitted) sample uses are legitimate or desirable, and which are not. This evaluation is specific for every single protocol and is part of the specific assessment to be made by the ethics committee examining the research proposal.
However, it is appropriate to consider the uses to which anonymized non-linkable samples should be limited. Some authors contend that, since such a sample is no longer traceable to its donor (and there are therefore no risks for the person concerned), it is legitimate, subject to prior informed consent and ethics committee approval, to use this sample for any legal purpose or at least for a broad range of purposes (for instance, "study of genetic influence on complex diseases or drug response").
It must also be considered who is allowed to work on the samples. The European Society of Human Genetics document "Data storage and DNA banking for biomedical research: technical, social and ethical issues"[27] underlines that freedom of movement of data and samples should be encouraged, subject to data protection.
To apply the most innovative technologies and skills to genetic research, it is often necessary or appropriate to send samples to other laboratories than those originally involved in the research, for instance to a specific laboratory where a particular analysis can be performed. Collaboration among different research groups appears essential to pursue the objectives of research: a research group could, for example, obtain important information from analysis of samples collected by another group working on the same disease. This exchange of samples among different laboratories is often difficult to predict and thus not usually specified in detail in the research protocol. The subject who supplied the sample may thus not know exactly where the sample is at a given time and who is using it.

The working group thinks that, when it is necessary or appropriate to transfer samples collected from one laboratory to another, and that such collaborations fall within the scope of the research, this must be subject to the following conditions:

- **the purpose is research, with no direct economic profit from the samples, that is to say the samples cannot be traded or sold for money. The activities performed on the sample can be paid for – for example, the promoter of a research can have a certain analysis done by a laboratory for payment**
- **sample use is in any case subject to the aims and conditions authorized by the ethics committee**
- **conditions and standards for sample safety, storage and protection must be exactly as stated in the study documents (see *"Storage"*)**
- **data protection methods and standards must be exactly as stated in the study documents**
- **the research promoter continues to be responsible for the collected samples**

- **the informed consent has to state the possibility that samples might be sent to other laboratories or research groups.**

Summing up and recommendations

It is useful to distinguish human tissue banks (biobanks) from collections of human biological samples generated within genetic research for the sole aims of the research concerned.
The biobank differs from the simple collection of biological samples within a genetic research setting in that the biobank has broader aims. In particular, one of the main purposes of the biobank is to make biological samples available to those who apply for permission to analyze them.

This distinction has important practical implications for issues such as duration of sample storage or how precisely the uses allowed must be described in the two cases.

The research promoter has the responsibility to guarantee that storage, analysis and transfer of samples will meet standards ensuring integrity, safety and control over the use of samples.

The level of identification of biological samples and data collected or produced by the study is a measure of how directly the name of the donor can be linked to them. The levels of identification indicated below are (i) those in the EMEA document, also adopted in the CPMP "Position Paper on Terminology in Pharmacogenetics" [(11)]; (ii) those adopted by the most recent document of the European Council Steering Committee for Bioethics (CDBI), "Recommendation of the Committee of Ministers to member states on research on biological material of human origin" [(3)]. The proper level of sample identification must be chosen with specific reference to the protocol to which it will be applied. It must be borne in mind that anonymous or anonymized samples and data (specifically non-linkable anonymized samples) do not make it possible to link sample or data with the subject's identity. This means optimal data protection, but by the same token the subject will not be allowed to withdraw consent, to request destruction of samples or receive results of sample analysis.

Storage and handling of samples after the termination of the study are proper and recommended if it is likely that stored samples will generate further scientifically useful information. In such cases, accurate information has to be provided, including storage methods, most likely use of the sample and possibility that samples may be sent to other laboratories/research groups.

At any time and without restriction, the subject must be allowed to demand destruction of the supplied sample, unless it is anonymized and non-linkable.

Biological samples collected in research can also be put to future uses if these are included in the aims of the original study, without need for renewed authorization and consent.

Samples can also be used for a purpose not included in the aims of the initial study if this purpose has been previously described in the protocol approved by the ethics committee and the patient has consented to it. In any other case, prior ethics committee approval and informed consent are needed (except for anonymized non-linkable samples).

The samples can be sent to other research groups or laboratories, on condition that this is done solely for purposes of scientific collaboration so that the sample can be subjected to certain procedures (for example, specific analyses); these can also be paid for, on condition that the guarantees regarding sample storage, possible uses and the level of data protection are exactly as specified in the informed consent. In such cases, this possibility must be mentioned in the informed consent and the required quality standards in handling and analyzing samples, data protection and

compliance with uses specified in the informed consent must be guaranteed. The samples cannot be sold for a monetary consideration.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

- **appropriate information and guarantees must be given regarding storage of samples and their possible use, particularly after the termination of the study**
- **the level of identification of samples and data must be specified and must be appropriate to the aims and methods of the specific study**
- **it must be guaranteed that samples are used only for future purposes described in the protocol and informed consent**
- **in any other case renewed ethics committee approval and informed consent must be obtained beforehand where anonymized non-linkable samples are not involved**
- **samples can be sent to other research groups or laboratories, on condition that this is done solely for purposes of scientific collaboration so that the sample can be subjected to certain procedures (for example, specific analyses); these can also be paid for, on condition that the guarantees regarding sample storage, possible uses and the level of data protection are exactly as specified in the informed consent. In such cases, this possibility must be mentioned in the informed consent and the required quality standards in handling and analyzing samples, data protection and compliance with uses specified in the informed consent must be guaranteed. The samples cannot be sold for a monetary consideration.**

Genetic analysis: which level of information

The most common approach for genetic analysis of samples is the "candidate genes" strategy, focusing on genes that can reasonably be involved in the phenomenon under investigation. There are also other techniques, like genome analysis with maps of SNPs (single nucleotide polymorphisms) and gene expression analysis through gene chips (or microarrays). For further information, see the text-box **"Candidate genes, genome scan and microarrays**".

One point to clarify is the level of information to be given regarding genetic analysis: is it necessary to specify which genes and polymorphisms are studied? To what extent is this consistent with the rapid evolution of genetic research?

The working group believes that the protocol has to specify the type of analysis that will be performed and provide detailed information about methods.

For example:

- *Qualitative and quantitative gene expression pattern (mRNA) via microarrays, including genes thought to be involved in the pathogenic mechanism of X, e.g. the genes involved in the signalling pathway of… or coding for…*
- *Sequence variations (DNA) in candidate genes thought to be associated with the response to Y, e.g. metabolism or target genes, or those involved in side effects in response to drug Y.*
- *Genome scan using polymorphic markers, e.g. SNPs.*

The working group believes that it is not necessary to specify in the protocol the actual genes or polymorphisms that are going to be analyzed or the detailed analysis methods, but it is sufficient to indicate the categories of genes analyzed (or, when using RNA, its origin – i.e. a precise indication of the tissue and the conditions under which the biological sample was stored before its extraction) and type of analysis, as indicated in the above examples.

Considering the rapid progress in discovering new genes and polymorphisms, and the difficulties in obtaining consent for new uses of the sample, study approval restricting the analysis to a specific list of genes would be a massive constraint to research without providing any real protection to the subject.

It is however appropriate, whenever possible, to give some examples of genes that could be analyzed, specifying that other genes could be used should their analysis provide an interesting scientific opportunity.

There is also the particular case of studies whose purpose is to investigate specific variations in specific genes. In this case it is necessary to list genes and polymorphisms, this level of detail being an integral part of describing the aim of the study.

Candidate genes, genome scan and microarrays

Analysis of candidate genes requires prior knowledge of which genes have a probable, if not already proven, involvement in the phenomenon under investigation. Specific polymorphisms are normally analyzed in genes related to the development of the investigated disease (or to related diseases); in pharmacogenetic studies, the analysis focuses on genes related to the absorption, distribution, metabolism and elimination of the drug, to its target or its side effects.

The analysis of the whole genome (whole genome genotyping) with high-density SNP maps is a very promising method, allowing the entire genome to be closely examined without requiring advance knowledge of any gene potentially involved in the phenomenon under study. This method could make the study of candidate genes obsolete, but to date it is still not available for routine use because it is costly and certain technological problems have still to be addressed. However, it is less expensive and simpler to apply this method to specific candidate genes or to limited regions of the genome, where a series of polymorphisms rather than of a single specific polymorphism are to be

analyzed. Since it allows analysis of far more positions within the genetic target region and determination of genetic haplotypes in extensive regions, this method is more informative than the analysis of specific polymorphisms. Since October 2005, when the first phase of the HapMap project was concluded, chips containing 500,000 SNPs which cover the whole human genome have been available. With SNP chips it is possible to determine genetic haplotypes within the whole human genome, and hence to identify the relevant gene or genetic variation.

Microarrays or gene chips are made up of a glass support of less than 2 square centimetres, subdivided into areas of 0.0005 square centimetres. On each of these areas it is possible to place about 100 single-stranded DNA sequences of about 20 base-pairs. Thousands of different sequences can be created, specific for single genes. For instance, there are already gene chips containing specific sequences for selective recognition all human genes (totalling about 27,000). From the tissues obtained in genomic research (for example, portions of neoplastic tissue), all the RNAs can be obtained and, once converted into cDNA (the single-stranded DNA transcribed from RNA), can be placed on the gene chip. Signals generated with specific fluorescence techniques differ in colour according to whether the cDNA hybridizes (because it finds a complementary sequence) or not (because it finds none), and in relation to the quantity of cDNA which binds. With this technique it is possible to measure changes in genetic regulation, for example as a result of disease, using rapid computerized analyses. The gene chip system allows analysis and comparison of the sequence and expression of thousands of genes. The fundamental difference compared with other methods is that the gene chip allows systematic analysis of large numbers of genes at the same time, making the familiar technique based on one-by-one analysis of single genes obsolete.

Summing up and Recommendations

DNA, RNA and protein analysis technology is a rapidly developing field that continues to offer new opportunities. Genetic research quickly discovers new genes, polymorphisms or associations of known genes and polymorphisms with different phenotypes. The requirement for prior selection and description of genes, polymorphisms or analytical techniques in a genetic research protocol could limit investigators' opportunities to apply new techniques or explore new scientific hypotheses, thus reducing the likelihood of generating useful results. Conversely, this level of detail does not increase the level of protection for the patient. The protocol must detail the genes and polymorphisms analyzed only in extremely focused studies which set out to analyze the genes/polymorphisms concerned. Otherwise, indications must be given about the types of genes to be studied and methods of analysis.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

- **it is appropriate that the analysis of the collected samples should benefit from all the instruments made available by scientific and technological advances, to increase the likelihood of the research giving useful results**
- **a research protocol rigidly restricted to the analysis of specific genes and polymorphisms might make it impossible to take advantage of such opportunities, while at the same time not providing greater safeguards for the patient**
- **the protocol must detail the genes and polymorphisms analyzed only in extremely focused studies which set out to analyze the genes/polymorphisms concerned**
- **For research with broader aims it is not necessary to specify in the protocol the genes and polymorphisms to be studied and the analysis techniques; indications about the types of genes to be studied and methods of analysis must nevertheless be given. It is appropriate to mention some examples if available.**

Study results and accessibility

A fundamental issue to be clarified in genetic research is access to the overall and individual results it produces.

Individual results

In recent years there has been discussion about whether to make individual results known to the subject, even if it is often not enviageable that the results will be clinically useful for the subject. In Italy the tendency has always been to allow the subject to choose whether to be informed, irrespective of how useful the data may prove. In Europe discussion on this issue is no longer relevant: in compliance with the principle stated in the European Council Convention on Human Rights and Biomedicine (art. 10), the Universal Declaration on the Human Genome (art. 5 *c*) and the UNESCO International Declaration on Genetic Data (art. 10), a subject who has participated in a research study has the right to be informed of individual genetic results and their implications if s/he so wishes, as well as of medical and scientific research results in cases where individual genetic and proteomic data or biological samples from which the data are extracted are used to such purposes. In addition, the European Directive on Personal Data Protection, states the subject's right to access "personal data" such as individual results originating from research. This means that:

- the subject has the right to access all personal data
- data must be given to the subject in a form comprehensible even for the "person in the street"
- there is no specific obligation to disclose the result to the subject without being asked to do so
- personal data revealing health status can be made known to the subject by healthcare professionals and health authorities, only through a physician named by the subject or by the study promoter.

Current Italian legislation on data protection, as stated in Law 196 of 30th June 2003, has implemented the same principle regarding right of access.

An important question is the way this information is given to the patient, and particularly whether there is the need to do so in the framework of a genetic counselling programme. **The working group agrees with the distinction between *Genetic Counselling* and *Genetic Information* as stated in the European Commission document "25 recommendations on the ethical, legal and social implications of genetic testing"** [(12)]. Counselling is restricted to diagnostic and pre-symptomatic testing for serious diseases and must be performed by professionals who have received specific training, i.e. experts in medical genetics. In the other cases information on genetic tests (including results) is part of the information that the physician gives the patient and must be given by physicians who have received suitable training.

Individual results of genetic research have therefore to be disclosed by the study physician in a readily understandable way to subjects who ask for them, irrespective of whether they might potentially be useful for the subject's health. The informed consent must clarify the nature of such results and state whether they may be useful for the subject's health, so that the s/he has information on the basis of which to decide whether to ask for the results.

In which circumstances does automatic disclosure of results to the patient, general practitioner or study physician become mandatory? The criterion to be followed when answering this question is the utility of the genetic results for the subject's health. Currently most genetic research consists of exploratory studies, so that there is little likelihood that results will be of immediate clinical use for subjects. In this case they are disclosed to the subject if s/he asks for them and are not to be disclosed to the study physician, subject's general practitioner or others. Individual results can thus be disclosed only to the subject on explicit request, through the study physician.

However, the first non-exploratory pharmacogenetic studies are now in progress and may generate results potentially useful for the health of the subject. Standards must therefore be set for assessing the level of scientific importance and clinical utility which make automatic disclosure of results to the patient mandatory. It is also crucial to address issues such as whether to disclose to the patient data that s/he asked not to be informed of, according to Article 10 of the Oviedo Convention. Criteria must also be established for assessing the circumstances in which it is appropriate that the subject's individual results are automatically disclosed to the study physician or subject's general practitioner, and for determining the subject's rights to limit this access. The working group acknowledges that at the time of writing there is not enough experience of this kind of study to allow detailed recommendations about the most appropriate approach to these issues. **As a general criterion, the working group recommends that in a study which may provide useful results for the subject's health these be given to the study physician so that s/he can take them into due consideration and share them with the patient. The informed consent must in this case specify whether results useful for the individual are expected and, if so, indicate that they will be disclosed to the study physician.**

Unexpected knowledge and collateral information

Collateral information is information created by genetic research but which is not part of its objectives. The possibility of this occurring is readily understandable, the axiom "one gene - one protein" now being totally obsolete – a gene is known to code on average for the production of about ten different proteins. It is also known that these proteins can be related to different phenomena or diseases (allelic heterogeneity or allelic series). For example, the known association of the gene APOE with Alzheimer disease was identified during research on the polymorphisms of familial hyperlipoproteinemia. The analysis of this gene can thus give information on the subject's susceptibility to both cardiovascular disease and Alzheimer. Another example is seen when study of the gene coding for a drug target in a pharmacogenetic study offers information about genetic influence on the progression of the disease, if the gene is related to both. Currently research may produce collateral information in only a few cases. In future, as knowledge of the different functions of proteins advances, more collateral information could emerge.
Unexpected knowledge (or data) is information that was not expected to emerge from the study. A typical example is the discovery of a "false paternity", which occurs quite often when clinical genetic testing of family members for a monogenic disease leads to the chance discovery that one individual is not the biological father of another. In most genetic research this cannot occur, either because single subjects are recruited and analyzed or because complex traits are examined. It must be borne in mind that the genome is not an "open book", and that in the great majority of cases the analysis highlights only the information actually sought.
The type of information produced by genetic research is therefore not always confined to the purposes of the study. However, as long as studies do not provide information with diagnostic content for a certain disease, there is only a minimal risk of accidentally discovering undesired medical information that might prove damaging for the subject.
It is rare for genetic research to produce unexpected knowledge or collateral information. This issue has to be managed only when this information is important for the health of the subject. It is neither practicable nor appropriate that the informed consent should include the choice as to whether any unexpected knowledge or collateral information should be made known to the subject, since the rarity and unpredictability of such events does not allow correct prior information enabling the patient to make a fully informed choice.

The working group thinks that the generation of unexpected or collateral knowledge is an issue to be managed only when this information is important for the health of the subject. With few exceptions, this is currently not the case in genetic research. The working group believes that it is neither practicable nor appropriate for the informed consent to include the

choice as to whether any unexpected knowledge or collateral information should be made known to the subject, since the rarity and unpredictability of such events does not allow correct prior information enabling the patient to make a fully informed choice. For example, a patient might wish to know collateral information regarding susceptibility to cardiovascular disease but not to tumours, but it is usually impossible to make such distinction beforehand.

Overall results

The production and dissemination of the results, as for any clinical research, is a duty required by the Helsinki 2000 document on use of information by the study promoter (publication/non-publication of positive/negative results).
Ethics committees to which a protocol is submitted must give explicit consideration to dissemination of the results and, according to a Ministerial Decree of May 12th 2006 (Official Gazette of August 22nd 2006), must ascertain that the protocol states: "the right to dissemination and publication of results by the researchers who performed the study is guaranteed, respecting current provisions about confidentiality of sensitive data and patent protection, and there can be no constraint on dissemination and publication of the results by the sponsor (art. 5, paragraph 3 (c))".
Currently participation in genetic research offers no direct advantage to the subject concerned. It is usually thought that s/he participates with the will to contribute to the advancement of scientific knowledge regarding the disease s/he has. The informed consent sometimes mentions this possibility as one of the benefits of the study. Producing the results of the study and making them explicit is thus a required and in some cases explicitly stated advantage for the scientific community and to society, and the subject participating in the research has wished to contribute to this. The investigators participating in the research should receive a report of overall results. These should be distributed in collective form, i.e. not allowing identification of the subjects involved.

Summing up and recommendations

In compliance with the principle stated in the European Council Convention on Human Rights and Biomedicine (art. 10), the Universal Declaration on the Human Genome (art. 5 *c*) and the UNESCO International Declaration on Genetic Data (art. 10), and consistent with the provisions of the European Directive on Personal Data Protection, a subject who has participated in a research study has the right to receive her/his individual genetic results if s/he so wishes, irrespective of their clinical utility. The informed consent must clarify the nature of such results and state whether they may be useful for the subject's health, so that the s/he has information on the basis of which to decide whether to ask for the results.

Genetic research might generate collateral information, i.e. information outside the study's objectives, or unexpected knowledge, i.e. information that was not expected to emerge from the study. This possibility is at present uncommon and usually the analysis reveals only what has been sought.

The European Commission document "25 recommendations on the ethical, legal and social implications of genetic testing" [(12)] makes a distinction between genetic counselling and genetic information and recommends that, with the exception of diagnostic and pre-symptomatic tests for severe diseases, the information generated by a genetic test can be disclosed to the patient by physicians who are not specialists in medical genetics but have received suitable training.

Individual results which have no immediate clinical utility must not be disclosed to any person other than upon request to the subject concerned, through the physician involved in the study. Individual results which may be useful for the health of the subject must be given to the study physician so that s/he can take them into due consideration and share them with the patient. The

informed consent has to specify whether any results are expected to be useful for the subject's health and, if so, that the study physician will be duly informed of them.

As recommended in the Helsinki 2000 document on use of information by the study promoter (publication/non-publication of positive/negative results), and as for any other biomedical research, the global results of genetic research must be made public.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

- **individual results of genetic research must be disclosed to the patient who asks for them by the study physician, irrespective of their clinical utility**
- **the informed consent must specify the nature of such results and whether they might be useful for the individual, so that s/he can make an informed choice as to whether to ask for them**
- **individual results which may be useful for safeguarding the health of the subject should be given to the study physician so that s/he can take them into due consideration and share them with the patient. The informed consent must specify whether results useful for the individual are expected and, if so, indicate that they will be disclosed to the study physician**
- **in many cases the research does not produce unexpected data or collateral information. This is an issue to be managed only when such information is important for the subject's health, but this rarely occurs in genetic research**
- **it is neither practicable nor appropriate for the informed consent to include the choice as to whether any unexpected knowledge or collateral information should be made known to the subject, since the rarity and unpredictability of such events does not allow correct prior information enabling the patient to make a fully informed choice**
- **production and dissemination of results, as for any clinical research, is a duty in compliance with the Helsinki 2000 document, in respect of the subject's wish to contribute to the advancement of scientific knowledge through participation in the research**
- **investigators who participate in the research should receive a report of overall results.**

Commercial and patent rights

The intense debate of some years ago concerning the appropriateness of patenting "biotechnological inventions" has in recent years lost momentum. The discussion was triggered by the evident opposition between the public Human Genome Project approach, which made all sequences freely available on the Internet, and Celera Genomics, which was developing its own private project and selling the sequences. Today Celera Genomics also makes sequences available free of charge, and other projects for the acquisition of fundamental scientific knowledge such as SNP and haplotype mapping, in which both public and private organizations are involved, make the information generally available on the Internet [29, 30].

The basic issue seems to have lost urgency – indeed, the European Directive on juridical safeguarding of biotechnological inventions [13], after considerable discussion, has yet to be implemented by almost all member states. In Italy this Directive has been only recently implemented, in Law n° 78 of 22nd February 2006, "Conversion into law, with modifications, of the Act of 10th January 2006, n. 3, implementing Directive 98/44/CE on juridical safeguarding of biotechnological inventions" [31]. The Law was published in issue n° 58 of the Official Gazette, dated 10th March 2006.

The most controversial aspect concerns the possibility, recognized by the Directive, of patenting biological material – and thus a DNA sequence – provided that it is isolated from its native environment or produced through a technical procedure, even if pre-existing in the natural state, on condition that its utility is described. For example, Incyte succeeded in patenting an isolated and purified form of the gene coding for the histamine receptor, but it would not have been able to patent the gene as such. One view of such matters is that sequences existing in nature should not be patented, even if they have been processed by human intervention.

The patent obtained by Myriad on the test for breast cancer susceptibility genes, BRCA1 and BRCA2, aroused considerable debate and criticisms. For further information about this topic, see the text-box ***"Patents on hereditary breast cancer genes"***.

Patents on hereditary breast cancer genes

The patents filed by the American company Myriad and the University of Utah Research Foundation on tests for the hereditary breast cancer susceptibility genes, BRCA1 and 2, aroused considerable debate and criticisms

Myriad identified these two genes in 1994 and 1995 respectively, after considerable knowledge on the subject had been made available by academic research. Between 2001 and 2003 the European Patent Office (EPO) granted Myriad different patents on these genes. Many prestigious institutions, included the Italian Society of Human Genetics, have firmly opposed the granting of these patents, putting strong pressure on the EPO to cancel them.

Following this strong opposition, the patents granted on BRCA1 were cancelled or limited in 2004 and in 2005, and today are not considered a threat to the free diagnostic use of the test in Europe. However, the question is not definitively resolved and Myriad has appealed against the decisions concerned.

In June 2005 the EPO granted the University of Utah Research Foundation the patent for the 697delT mutation of the BRCA2 gene, analyzed for "diagnosing a predisposition to breast cancer in Ashkenazi Jewish women in vitro."

This grant aroused considerable criticism. The European Society of Human Genetics, which is opposed to this patent, believes that it could be discriminatory against the Ashkenazi Jewish ethnic group. A physician who intends to perform the test has to ask a female patient if she is an Ashkenazi Jew. If she is, she will probably have to pay for the test because the health system has in turn to pay patent royalties to the University of Utah Research Foundation. If she is not, or she simply states that she is not, the test can be given free bacause the health system does not have to pay for any license.

The criticisms against Myriad did not concern only the patents, but above all the fact that the firm has exclusive rights on the analysis of these genes.
Whoever wants to perform these analyses has to send a tissue sample to Myriad, which charges a fee for the analysis and the result. No other laboratory can perform the analysis. An alternative that would certainly have sparked less controversy would have been patenting the test on these genes and allowing other laboratories to perform it for a fee. This would be consistent with the recommendation of the Nuffield Bioethics Council about patents on DNA sequences for purposes of diagnostic testing (see the text-box ***"The Nuffield Bioethics Council and DNA patents"***).

Information about the prerequisites for filing a patent on a biotechnological invention can be found in the text-box ***"Genes and patents"***.

Genes and patents

What is a patent?

The patent is a form of intellectual propriety on an invention. It confers the patent owner exclusive use of the invention. A legal monopoly of this kind is limited in scope, duration and territorial extension. It is valid only in the country or countries in which it has been applied for and granted (there is still no such thing as a worldwide patent), it generally lasts 20 years from when the application has been made (after which anyone can use the invention) and only for the uses indicated in the patent application, thus delimiting the "borders" of the patent monopoly.
The patent is issued by the state on the inventor's application and in practice officializes an agreement to protect the inventor and at the same time to favour research and innovation, since:

- the state guarantees to the inventor the right of exclusive use of the invention, or the right to exclude others from making, using or selling the invention, or drawing any profit from it without the patent owner's authorization.
- the inventor is obliged to make the invention public. In the published patent, it must be stated in detail what has been invented, what the invention is for and how it has been obtained, with all the necessary instructions to reproduce and use it.

What is patentable?

Not all inventions are patentable. To be patented an invention must:

- be either a product or a process or a new application of an already known product
- satisfy the following criteria:

- novelty: it cannot be already known (for example, it cannot have already been published)
- non-obviousness: the invention must not be obvious to a person with ordinary skill in the field to which the invention belongs
- industrial applicability or utility: it must in some way be applicable or useful to industry.

Innovations such as discoveries, scientific theories, mathematical methods or methods for surgical or therapeutic treatment of the body cannot be patented. It is also not possible to patent inventions that are contrary to law and order or to morality. There are some differences between countries and between Europe and the United States – for example, in the U.S.A. methods for treating the human or animals are patentable.
To simplify granting of patents in the different European states, the European Patent was introduced.
The European Patent is issued by the European Patent Office (EPO), based in Munich, following a single application filed in just one language (English, French or German) and prior verification that patentability criteria are fulfilled.
Once granted, the patent is subdivided into a number of national patents.
The advantage of the European procedure is that a single application is filed and a single patent granted (the text and claims thus remaining unchanged), valid for all member states of the European Patent Organization designated by the applicant.

The patent in the biotechnology field

In legal terms the fundamental patentability criteria can be applied indiscriminately to any invention. Patenting a new drug obtained after chemical synthesis is no different from patenting a drug achieved with recombinant DNA technique or a genetic test: in each case it is necessary to demonstrate that the previously listed criteria are satisfied. This principle has been confirmed by European Directive 98/44 on the patentability of biotechnological inventions, which specifies how general patentability criteria are applied to biotechnological inventions [(13)].

The Directive confirms that biotechnological inventions contrary to law and order and to morality (such as human cloning procedures, modifications of the human stem cell line, use of embryos for commercial purposes) cannot be patented. The Directive also confirms that the patent criteria in use for all patentable inventions also apply to genes/DNA sequences.

The simple discovery of a partial or complete gene sequence is not patentable, in compliance with the general criterion stating that a discovery as such cannot be patented. On the contrary, it is possible to patent genes/DNA sequences isolated from the human body or produced through a technical procedure with innovative methods. To satisfy the industrial applicability/utility criterion, when patenting a DNA sequence the patent application must describe the function of the sequence concretely. For example, in the case of a sequence coding for a protein the application must state "DNA sequence coding for protein X, having function Y ". A DNA sequence that does not have a clear function is therefore not patentable. ESTs or SNPs, used as probes or markers, are not patentable since they have no utility as such, but are simply a research tool. They could be patented if it were possible to describe their application in a certain disease.

The purposes of the Directive were to clarify a number of issues, which were at the time extremely controversial, regarding patentability of genes/DNA sequences and to harmonise the legislation of the different member states. The Directive has actually aroused further discussion and has only recently been implemented by most member states, with the exception of Luxemburg and Latvia. In the countries that did not issue the Directive, the European Patent Convention, recognizing the patentability of gene sequences, still applies.

There used to be some differences in sequence patentability between Europe and the United States, the requirement that the utility of the invention be to demonstrated being less binding in the USA. The review of the US Patent and Trademark Office guidelines on the utility criterion has to all intents and purposes brought US and European criteria on biotechnological inventions into line.

Implementation of the European Directive on patentability of biotechnological inventions in Italy: Law 78 of 22nd February 2006 [(31)]

Italian law has implemented, without substantial modifications, the European regulations on the patentability of biotechnological inventions. It is possible to patent the following inventions, subject to fulfilment of the novelty, non-obviousness and industrial utility criteria:

- a biological material, isolated from its native environment or produced through a technical proceeding, even if pre-existing in the natural state;
- a technical procedure through which a biological material is produced, modified or used, even if pre-existing in the natural state;
- any innovative application of an already patented biological material or technical procedure;
- an invention involving an element isolated from the human body or obtained by other means through a technical procedure, even if its structure is identical to that of a natural element, on condition that its function and industrial application are concretely indicated, described and specifically claimed. By technical procedure is meant what only a human being is able to perform and nature alone cannot achieve;
- an invention involving plants or animals, characterized by the expression of a determined gene and not by the whole genome, if the application is not limited in technical

terms to obtaining a specific plant variety or animal species and is not dependent on solely biological procedures, as stated in article 5, paragraph 6.

The following are not patentable:

- the human body, from the time of conception and during the different stages of development, as well as the mere discovery of one of the body's elements, including the complete or partial sequence of a gene, the aim being to guarantee that the patent right is exercised in respect of the fundamental human rights of dignity and integrity as well as of the environment;
- methods for surgical or therapeutic treatment and diagnostic methods, in humans or animals;
- inventions whose commercial use is contrary to human dignity, law and order, morality, healthcare, the safeguarding of the environment and of human or animal life, the preservation of plants and biodiversity, and the prevention of severe environmental damage, consistent with the principles stated in art. 27, paragraph 2 of the Agreement on trade-related issues in intellectual property rights (TRIPS);
- a simple DNA sequence or partial gene sequence used to produce a complete or partial protein, subject to proper indication and description of a function relevant to assessment of the industrial utility criterion and to a specific claim regarding this function;
- plant varieties and races of animal, as well as the "essentially biological" procedures to produce animals or plants;
- new plant varieties in which the invention consists solely in genetic modification of other plant varieties, even if such modification is produced by a genetic engineering procedure.

The working group acknowledges the importance of distinguishing what is being patented. A patent obtained by a public or private firm for an application such as a test is different from a patent concerning sequences or SNPs. For example, the patent obtained by a pharmaceutical firm on a pharmacogenetic test used for correct prescription of a drug developed by the firm itself is similar to the patents on the drug or on any diagnostic test. A patent on gene sequences or SNPs is much more disputable, even if it must not be forgotten that, as previously mentioned, sequences and SNPs are today widely and publicly available.

Any conclusions on issues such as which commercial and patent rights should be defined by a specific regulation is beyond the scope of the present guidelines on genetic research.

The working group believes that, if it is possible that research results generate commercial advantages for the research promoter and/or patents, this must be specified in the informed consent and it must be stated that there will be no economic advantage for the subject.

Every commercial or patent right applies to the results of research and not to the collected samples, to which it is not possible to apply the concept of property. Nobody can claim juridical ownership of collected samples, hence the need to define the rights and duties of research promoter, researcher and subject, and to describe them in the protocol and in the informed consent. For example, the research promoter has the duty to store samples correctly, protecting their integrity and guaranteeing data protection, and s/he has the right to use the results of the research both commercially and through patents. The subject has the right to request the destruction of samples at any time, but s/he has no right to enjoy the economic benefits stemming from the use of the study results. The document issued by the European Society of Human Genetics, "Data storage and DNA banking for biomedical research: technical, social and ethical issues" [(27)], states that the subject must be given "primary control" of DNA and data, while the researcher or the person who analyzes data and samples is the "guardian".

The Group upholds the tenets stated in many documents, including the joint Italian Society of Human Genetics - Telethon guidelines on Genetic Biobanks [(26)] and the most recent document of the European Council Steering Committee for Bioethics (CDBI), "Recommendation of the Committee of Ministers to member states on research on biological material of human origin"[(3)],

regarding the need to ensure that no direct profit can be made from samples – i.e. to prevent their commercialization or sale.
The Nuffield Council on Bioethics publication "The ethics of patenting DNA" [32] provides a very detailed discussion on ethical aspects of patentability of DNA. See the text-box ***"The Nuffield Bioethics Council and DNA patents".***

The Nuffield Bioethics Council and DNA patents

The Nuffield Bioethics Council stated its opinion on DNA patents in the document "The ethics of patenting DNA". Among the central issues raised are the following considerations:

- the patent system offers great benefits to society
- it is appropriate to wonder whether the application of such a system to DNA sequences really contributes to the aims of stimulating innovation for public benefit and recompensing individuals who produce useful new inventions
- DNA sequences are essentially genetic information and cannot be considered in the same way as a chemical for the purposes of the patent system: the debate on the possibility to patent DNA sequences should be reviewed accordingly
- the novelty, non-obviousness and utility criteria requested for the patentability of DNA sequences have not been sufficiently applied yet and many already granted patents are of doubtful validity.

There are 4 different applications for a DNA sequence, and in this respect the Nuffield Bioethics Council gives different recommendations about the appropriateness of granting patents:

- **Diagnostic tests**: for these tests it is suitable to apply the existing criteria rigorously, with specific reference to the "non-obviousness" crierion, and to consider the possibility of making it mandatory to grant user licenses for tests based on genes
- **Research tools**: this kind of patent should be deterred, by strict application of the utility criterion
- **Gene Therapy**: this kind of patent should be granted only in rare instances, since using an already known gene for gene therapy does not satisfy the non-obviousness criterion
- **Therapeutic Proteins**: when a patent application is made for a DNA sequence that will be used as a therapeutic protein, it should be limited solely to the specific protein described.

Summing up and recommendations

Patents on biotechnological inventions are based on the same requirements as any other invention: novelty, non-obviousness and utility. European Directive 98/44 (implemented in Italy as Law n° 78 of 22nd February 2006) [13] also specifies that neither the human body nor the mere discovery of one of its parts is patentable, while a biological material isolated from its native environment or produced through a technical procedure, even if pre-existing in the natural state, is. This principle applies to the complete or partial sequence of a gene: the mere discovery of a gene sequence is not patentable, but the sequence becomes patentable if it has been isolated from the human body or produced through a technical procedure involving innovative processes, and if there is a description of a possible application for that sequence.

Regarding the possibility that the use of patented tests or sequences may lead to a situation of monopoly, a distinction must be made between patents and licenses. In the light of this distinction, one view of the matter is that monopoly situations arise as a result of licenses not being granted rather than of patents being registered.

The results of genetic research may create potential for commercial or patent exploitation. The subject who agrees to participate in the research must be aware of this possibility.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

- **samples collected within a study cannot be sold or bought for profit**
- **however, the research results may create commercial advantages and/or patents for the research promoter**
- **when a patent or commercial exploitation is envisageable, this must be stated in the informed consent and it must be clarified that there will be no economic benefit for the subject.**

Insurance cover

The potential risks of genetic research include the physical risks typical of any research and "non-physical" or "information-related" risks, for example the moral and/or material damage that can occur from improper use of a subject's genetic information by other persons (employer, insurance companies, or even relatives). (See ***"Non-physical risk"***). Insurance policies **cover the reimbursement of any material damage caused by the research performed in accordance with the protocol**.
Physical risk is in practice the only insurable risk, and must be adequately covered by an insurance policy.
It is not possible to stipulate a policy covering risks of personal or moral damage, since insurers are currently unable to estimate the risk and thus calculate a premium.
On the other hand, if a study is correctly planned and performed, the concrete likelihood of the subject sustaining any non-material damage is very limited. See the text-box "***Measures to limit risks related to release of information".***

Measures to limit the risks related to release of information

The following measures help minimise the risk of non-physical damage as a result of information produced by genetic research:
• providing results of the genetic analysis solely to the subject if s/he asks for them
• informing the subject of the risks s/he can run by disclosing the results.
• using coded (or anonymized) samples and data from which only the investigator (if anybody) can trace the subject's identity
• preserving the samples in monitored rooms and limiting access to research staff
• documenting every activity performed on the samples: deliveries to other laboratories, destruction etc.
• adequately protecting the paper or electronic database from external intrusion and limiting access solely to those working to research staff.

It is therefore of vital importance, to safeguard not only for the subject but also the investigator and the research promoter, that the study protocol and informed consent adequately clarify critical issues in terms of non-material risks and related liabilities.
If the subject suffers non-physical damage not covered by the insurance policy, this can cause litigation to be settled in court or by the National Data Protection Office. In such cases, everything written in the protocol, informed consent, study archives or records of the laboratory where samples are stored and analyzed can have a considerable bearing on the issue of liability.

Summing up and recommendations

In genetic research, physical risk is in practice the only insurable risk, and must be adequately covered by an insurance policy. It is currently not possible in most cases to stipulate a policy covering risks of non-physical damage (i.e. any personal or moral damage resulting from release of information collected or generated in the study), such damage being difficult to evaluate.

Correct planning and management of the research, in particular of samples and individual genetic data, helps minimise the risk of any non-physical damage as a result of genetic research.
If the subject suffers non-physical damage not covered by the insurance policy, this can cause litigation to be settled in court or by the National Data Protection Office. In such cases, everything written in the protocol, informed consent, study archives or records of the laboratory where samples are stored and analyzed can have a considerable bearing on the issue of liability.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

• an insurance policy is essential to cover risk of material damage

• to protect the subject (but also the investigator and the study promoter) from non-physical risks, it is important that the protocol and informed consent clearly state critical issues regarding non-material risks and related liabilities

• the most critical issues are storage, use and level of identification of samples and data, together with the disclosure of individual genetic analysis results and correct, exhaustive information to the patient about possible risks.

Genetic research on minors and mentally disordered subjects

European Directive 2001/20 on Clinical Research, implemented in Italy as Law 211/2003, gives specific consideration to research on minors and mentally disordered subjects, providing the basic norms for genetic research on these categories.

In general terms, the working group believes that, considering the complex, major implications raised by genetic research, the subject who gives consent should be able to fully assess advantages and disadvantages of participating in the study. This principle, as such, would preclude the possibility of genetic studies on minors or mentally disordered subjects.

The working group believes that the conditions specified by the Convention of Oviedo for clinical tests on minors and mentally disordered subjects can also be applied to genetic research.

See the text-box ***"Genetic research on minors and mentally disordered subjects: principles".***

Genetic research on minors and mentally disordered subjects: principles

- The importance of genetic information implies the faculty to take the responsibility for decision and thus demands maturity and awareness
- This requirement can be waived only to pursue the best interest of minors or mentally disordered subjects and with the explicit consent of the individual's legal representative
- When the result of the test does not allow any efficacious preventive treatment or improvement of the minor/mentally disordered subject's health, being submitted to the test is not in the best interests of the individual concerned
- In any case, the benefit concerned must have been previously demonstrated on adults or individuals in full possession of their faculties, except in the case of a disease typical of minors/mentally disordered subjects
- Pre-symptomatic testing (predicting future onset of a disease) should be postponed until the subject has reached adulthood and can thus make a fully autonomous decision, except when there are concrete possibilities of therapy or efficacious preventive treatments before adulthood and subject to explicit consent from the subject's legal representative.

Although recommending general caution when performing genetic studies on minors, the working group indicates some cases and conditions in which these genetic tests can be carried out:

- when the disease under investigation is typical of the minor or mentally disordered subject
- in other instances, when the advantage is proven and only after some adult studies
- with regard to pre-symptomatic testing (e.g. Hungtington disease), only if there are concrete possibilities of therapy or efficacious preventive treatment before adulthood. In all other cases the test must be postponed until adulthood
- when the test establishes within acceptable limits a prognosis of disease requiring clinical and/or pharmacological monitoring to delay onset or severity of symptoms (e.g. myotonic dystrophy).

It is necessary to obtain, whenever possible, the consent of the minor or the mentally disordered subject concerned in addition to that of a parent/guardian

The working group fully supports the authority of parents/guardians when deciding for minors/mentally disordered subjects in the above-mentioned cases, though recognizing the potential implications of deciding for minors/mentally disordered subjects (See the text-box "***The consequences of deciding for minors"***).

The consequences of deciding for minors

The parent/guardian who decides to authorise a genetic test on a minor exposes the subject concerned to potential consequences:

- violation of the minor's right to decide whether to perform the test once of age
- violation of the right to confidentiality of results
- potential damage on the minor's self-esteem, particularly if the test result is disadvantageous
- changes in the relationship between parents and a child in whom disease has been predicted: parents could become over-protective or discriminatory towards brothers and sisters
- discrimination of the minor at school and with regard to investment in education
- consequences on future job prospects and ability to build important, stable relationships.

The real likelihood that such consequences could arise deserves careful consideration.
If the minor undergoes a genetic test with clinical aims, that is to diagnose a genetic predisposition, there is the possibility of such consequences arising, particularly if the test concerns monogenic diseases or high penetrance genes.
This likelihood is considerably reduced if the test is performed within genetic research, not intended to diagnose but only to study a genetic phenomenon.
See also ***"Genetic testing for diagnostic purpose and research purposes"***.

Summing up and recommendations
European Directive 2001/20 on Clinical Research, implemented in Italy as Law 211/2003, gives specific consideration to research on minors and mentally disordered subjects, providing the basic norms for genetic research on these categories.

Since general caution is needed when performing genetic studies on minors and mentally disordered subjects, it is recommended that whenever possible consent be obtained from the subject concerned in addition to that of a parent/guardian.

Studies on minors or mentally disordered subjects are legitimate when the disease studied is typical of the minor or mentally disordered subject, or when there is an advantage for the minor/mentally disordered individual, and even then only after studies have been carried out in adult subjects.

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:
• studies on minors or mentally disordered subjects are legitimate only when the disease studied is typical of the minor or mentally disordered subject, or when there is an advantage for the minor/mentally disordered individual, and even then only after studies have been carried out in adults
• whenever possible consent must be obtained from the subject concerned in addition to that of a parent/guardian.

Protocol and informed consent

Any collection of human blood samples or other human tissues used for research involving genetic tests has to be under the control and supervision of the ethics committee, which focuses its examination particularly on the protocol and informed consent used for the study.
Any new study or novel use of biological samples not envisaged and authorized at the time of sample collection must be submitted to the ethics committee. If samples and data are not anonymized it will also be necessary to obtain new consent from the subject. When samples or data are anonymized it is not possible to ask for new informed consent and in such cases only the ethics committee can guarantee a responsible use of biological resources.
In general terms it can be stated that both the protocol and informed consent for genetic research must be drafted and assessed in compliance with the same general criteria as for any biomedical research protocol. In addition, these documents must contain specific information concerning the release of individual results of the genetic research (to whom they will be given and under which circumstances), as well as the collection, storage and use of the collected biological samples and data (see ***"Individual results"***, **"Biological samples: level of identification, storage and use"** and the check lists for the drafting and assessment of the protocol and informed consent in genetic research).
In particular, the following must be described:
• collection method, level of identification, duration of storage and future use of samples, guarantees of safe storage, possibility for the subject to ask for sample destruction, possibility of samples being sent to other laboratories
• duration, methods and safety measures for handling and storage of data collected or produced within the research, possibility of samples being sent to other laboratories and, if these are abroad, guarantee that there will be the same safety standards with regard to data protection
• authorization by the subject to use data/samples for further purposes (if relevant)
• possibility for the subject to obtain personal results on request, even if not useful for his/her health
• who can obtain individual results in addition to the subject concerned.

The US National Bioethics Advisory Commission (NBAC) document "Research involving Human biological materials: ethical issues and policy guidance"[33] recommends (recommendation 5) that the ethics committee ask the researcher for:
• thorough justification of the study design, included procedures to minimise risks
• a complete description of the sample collection process
• procedures for access to the subject's medical data
• description of the mechanism used to minimise risk of unnoticed release of confidential information.

The working group recommends that, when the genetic investigation is part of a clinical trial, specific informed consent must be obtained regarding the subject's involvement in the genetic testing. The subject should be allowed, if s/he so desires, to participate in the clinical study but not the genetic testing. It must be underlined that in some instances such a choice might not be possible. There are already ongoing pharmacogenetic studies in which patients are selected for different randomization groups according to their genotype. In such cases the genetic test is mandatory for participation in the clinical study and the patient cannot choose whether to accept or refuse it. Such studies may become more frequent in future, making it more difficult not only to distinguish clearly between clinical study and genetic study but also to give the patient a choice between agreeing to one or both of them.

With regard to information given about genetic research and the difficulty for many subjects to understand such topics, particular attention must be paid to the completeness and comprehensibility

of the information provided. Concerning the general formulation of the informed consent, this can be administered in various forms, according to the degree of freedom for the researcher and protection for the subject. See the text-box ***"Types of informed consent".***

Type	Characteristics	Pros and cons
"Restricted" model	• Detailed description of sample uses • Sample destroyed at the end of the study	Reduces potential risks but does not encourage research in a field which is promising for public health
"Extended" model	• Possibility of non-specified "other uses" of samples in addition to those described (for anonymized samples) • Sample(s) freely provided to other researchers • Sample(s) stored for an indefinite period	Supports the progress of research, but the subject is not involved in the further purposes to which the sample(s) provided will be put
"Intermediate" models	• Combines features of the above models according to the aims and implications of the research	Probably the best, tending to balance the progress of research and the duty to provide the subject with information

The National Committee for Bioethics document "From pharmacogenetics to pharmacogenomics" [(5)] points out the limitations of an over-restricted consent, which requires a laborious procedure for re-obtaining consent if opportunities for new analyses emerge.

Summing up and recommendations

Both the protocol and informed consent for genetic research must be drafted and assessed in compliance with the same general criteria as for any biomedical research protocol. These documents must also contain detailed information regarding disclosure of individual results from the genetic study and the collection, storage and use of any biological samples involved.

When the genetic investigation is simply part of a study with more extensive purposes it is appropriate that the subject should be free to choose whether to participate, without forfeiting the right to participate in the non-genetic part of the study. It must be underlined that in some instances such a choice might not be possible. There are already ongoing pharmacogenetic studies in which patients are selected for different randomization groups according to their genotype. In such cases the genetic test is mandatory for participation in the clinical study and the patient cannot choose whether to accept or refuse it. Such studies may become more frequent in future, making it more difficult not only to distinguish clearly between clinical study and genetic study but also to give the patient a choice between agreeing to one or both of them.

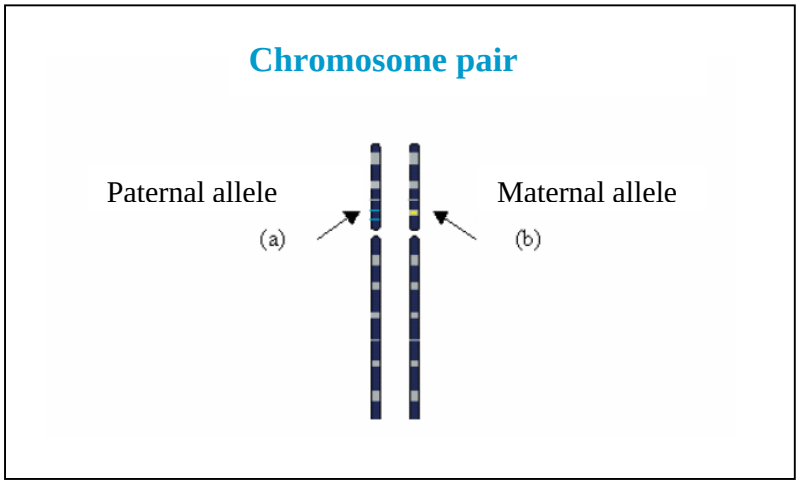
A good informed consent form should achieve optimum balance between the freedom of research and the level of detail in the information given to the subject. An indication of what should be contained in a genetic research protocol and informed consent is provided in the chapter **"Check-list for drafting and assessment of a genetic research protocol and informed consent".**

In drafting or assessing clinical research protocols in genetics, the working group recommends that the following points be borne in mind:

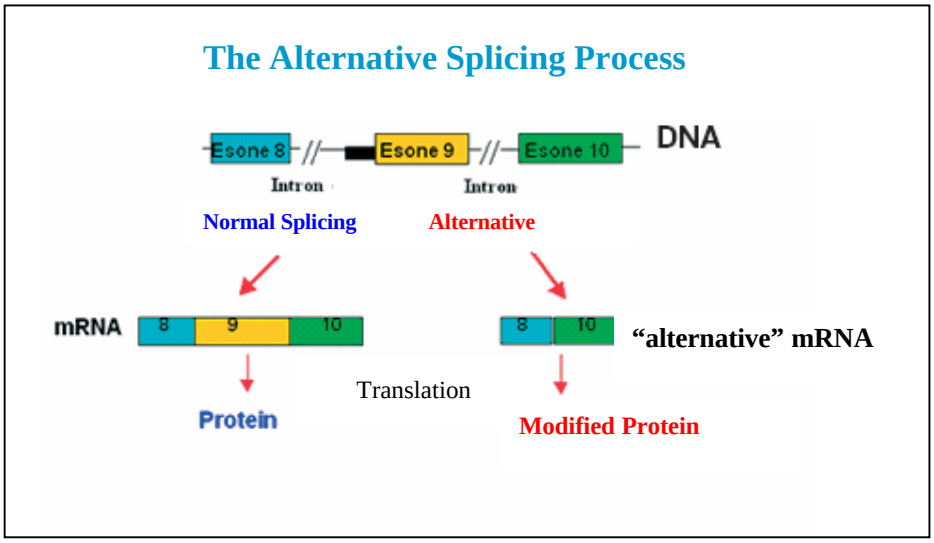
- **the protocol and informed consent for genetic research must be drafted and assessed in compliance with the same general criteria as for any biomedical research protocol. These documents must also contain detailed information regarding disclosure of individual results from the genetic study and the collection, storage and use of any biological samples involved**
- **when the genetic investigation is part of a clinical trial, specific informed consent must be obtained regarding the subject's involvement in the genetic testing. The subject should be allowed, if s/he so desires, to participate in the clinical study but not the genetic testing. It must be underlined that in some instances such a choice might not be possible. There are already ongoing pharmacogenetic studies in which patients are selected for different randomization groups according to their genotype. In such cases the genetic test is mandatory for participation in the clinical study and the patient cannot choose whether to accept or refuse it. Such studies may become more frequent in future, making it more difficult not only to distinguish clearly between clinical study and genetic study but also to give the patient a choice between agreeing to one or both of them**
- **the completeness and comprehensibility of information given to the subject are particularly important for genetic research**
- **the chapter "Check-list for drafting and assessment of a genetic research protocol and informed consent" indicates what must be included in a genetic research protocol and informed consent.**

GLOSSARY

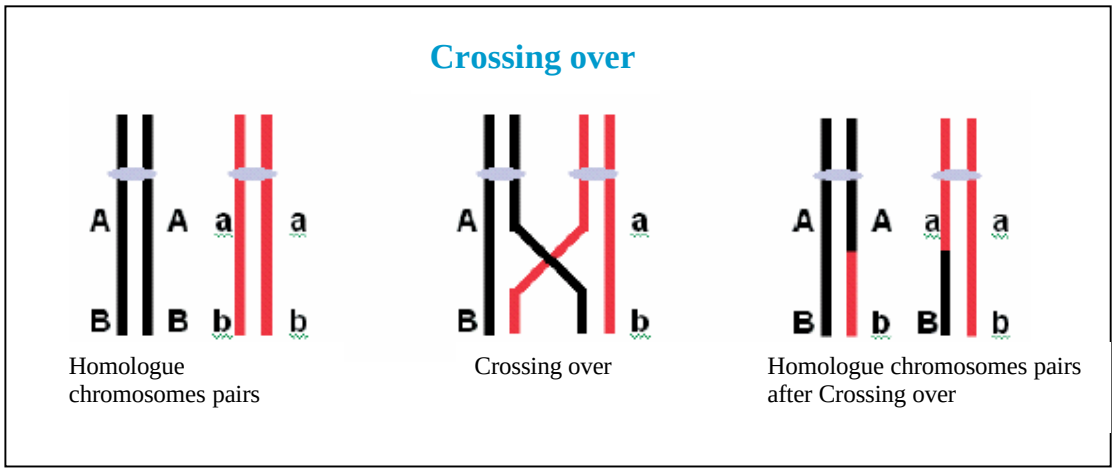
Adenine (A): a nitrogen purine nucleobase contained in DNA and RNA. In DNA adenine pairs with thymine.
Allele: one of the alternative forms of a gene in a specific chromosome localization (locus). Every individual has two copies of each gene, one inherited from the father and the other from the mother. The 2 copies of a gene, called alleles, can be identical (the organism is homozygous for that character) or different (heterozygous). The two alleles in a locus of a specific individual constitute a genotype.



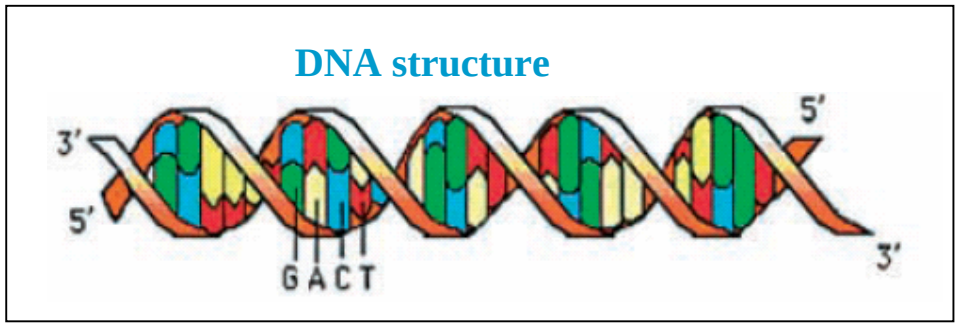
Alternative splicing: occurs by modification of the splicing process: the exons are re-bound with a different structure to that which is immediately deducible from the DNA sequence (in the example below, exon 9 was eliminated). This process gives rise to similar proteins differing in structure and usually in activity.



Amino acid: one of the 20 fundamental chemical units that can be linked in long chains to form polypeptides or proteins (i.e. alanine, methionine).
Anticipation: a phenomenon in which a genetic disease shows increasingly early onset and severity from one generation to the next.
Association study: investigation aimed at comparing the frequency of a specific allele between two populations – affected individuals versus control individuals, i.e. subjects free of disease. If an allele is more frequent in affected individuals, that gene could in some way be associated with the disease.
Autosomes: the chromosomes from 1 to 22 – i.e. all chromosomes other than the sexual ones, X and Y.
Bioinformatics: science dealing with information technology to analyze biological data. Needed to support the huge quantity of data and DNA sequences, mainly in genomic research.
Candidate gene: a gene whose function or position suggests it could be involved in the development of a disease or in the manifestation of a character.
Carrier: heterozygous individual for a mutant allele that generally causes a manifest phenotype (e.g. a disease) only in the homozygous condition.
cDNA (complementary DNA): a DNA strand synthesized in vitro from mRNA by reverse transcription, and thus devoid of the introns in genomic DNA. It is useful, for instance, in production of large amounts of a protein through cloning and gene expression.
cDNA collection: a collection of cDNA clones, representing all the transcripts (mRNA) of a specific tissue.
Centromere: a constriction of the chromosome separating the short arm from the long arm. Its main function is to guarantee the correct attachment of chromosomes to spindle fibres during cell division.
Character: a feature of an individual belonging to a certain species, for which there can be different types.
Chromatids: the two identical copies of a chromosome obtained after DNA duplication, bound by the centromere; during mitosis the chromatids separate and stay independent as newborn chromosomes.
Chromatin: the substance making up chromosomes, consisting of DNA, proteins and RNA.
Chromosomal mutation: any change in the chromosome number or structure (e.g. acquisition of another copy at chromosome 21 determines trisomy 21 or Down syndrome)
Chromosome: structure made of chromatin, which can be easily seen as a separate entity during cell division (see **mitosis**). When the cell is not dividing the single chromosomes are not distinguishable under the microscope, since the chromatin is relaxed and appears as a shapeless mass. Human beings have 46 chromosomes organized in 23 pairs. From 1 to 22 are the autosomes, while X and Y are sexual chromosomes.
Clone: all the cells deriving from the same progenitor cell, genetically identical to each other.
Cloned gene: a gene is defined as "cloned" when its sequence is reproduced in a laboratory.
Cloning: creation, "in vivo" in bacteria or "in vitro", through the PCR technique, of a vast number of identical DNA molecules. It is necessary to use enzymes that are able to cut, modify and bind different DNA segments.
Codon: three nucleotides specifying an amino acid or a stop signal in the synthesis. (e.g. GAG= glutamic acid)
Crossing over or recombination: exchange of DNA segments between homologous chromosomes by rupture and reconnection. This phenomenon occurs during meiosis (see **meiosis**).

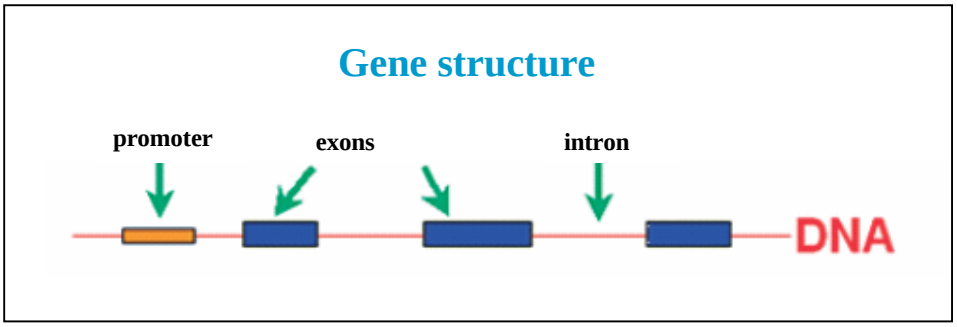


Cytogenetics: cytological approach to genetics, dealing mainly with the microscopic study of chromosomes.
Cytosine (C): a pyrimidine nitrogen nucleobase contained in both DNA and RNA. In DNA it pairs with guanine.
Degeneration: the characteristic redundancy of the genetic code, as a result of which some amino acids are encoded by more than one codon.
Denaturation: drastic change in the conformation of a protein or nucleic acid, such as transition of DNA from a double helix to single helix structure. Obtained with a temperature increase or chemical treatment with denaturing agents.
DGGE (Denaturing Gradient Gel Electrophoresis): this method consists in causing the migration of double-stranded DNA fragments through an increasing denaturing gradient (temperature or chemical), slowly separating the two helices. When the two strands separate, the migration stops. A change, even in just one base pair, may alter the position into which the mutated fragments migrate by comparison compared with normal fragments, thus identifying the mutation.
Diploid: cell or organism with a double chromosome set compared with that in gametes.
DNA (DeoxyriboNucleic Acid): initials for DeoxyriboNucleic Acid. A polymer made up of 4 different nucleotides, each composed by a nitrogen base (adenine, thymine, cytosine and guanine), a 5-carbon sugar (deoxyribose) and phosphate groups. Because of their chemical properties the nitrogen bases bind two by two. A double-helix structure thus forms, with the two polynucleotide chains running anti-parallel (in opposite directions).



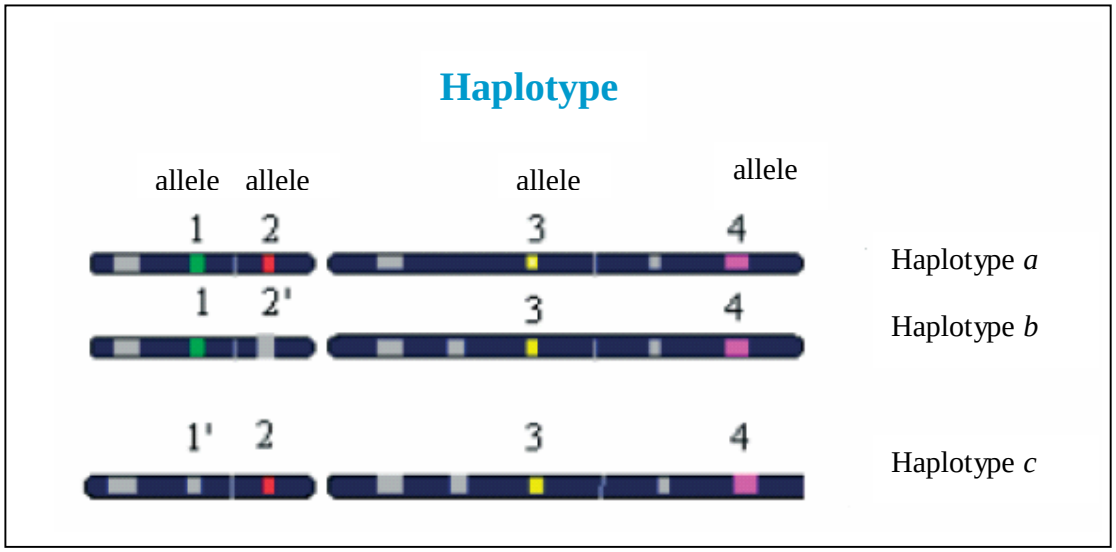
DNA polymerase: the enzyme operating semi-conservative DNA duplication. The double helix opens and any chain acts as a template for the synthesis of a new complementary chain.
Domain: portion of a protein having its own function. The combination of the different domains defines the global function of the protein.

Dominant: a character defined as "dominant" shows an effect even in heterozygotes (that is individuals with one normal allele and one mutated one).
Electrophoresis: technique based on the migration of a substance into a porous matrix (polyacrilamide or agarose). The electric field applied to a solution containing proteins or nucleic acids causes the molecules to migrate in a certain direction at a speed that depends on their size and mass. The molecules are thus separated and identified.
Epistasis: phenomenon in which a gene interferes with another and alters its phenotypic manifestation. For example, a clinical condition deriving from the effects of many genes may be different from the phenotype caused by the effect of a single gene, as a consequence of the interaction among different genes.
Exon: any coding gene sequence; it is transcripted into mRNA and translated into a protein. In the gene exons alternate with introns, which contain non-coding sequences that do not give origin to any protein.



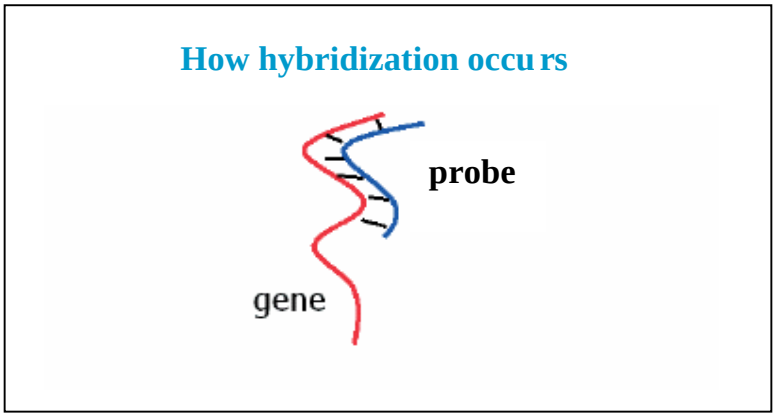
Expression: the process for transferring the information coded by the gene into a functional product, that is the protein. The term is used to indicate the type and quantity of proteins produced.
FISH (Fluorescence In Situ Hybridization): cytogenetic investigation using DNA fluorescent probes that can specifically bind to complementary sequences on chromosomes fixed on to glass slides and recognisable by UV microscopy.
Functional genomics: extensive study of the expression and function of multiple genes in biochemical processes of living organisms.
Gametes: female and male reproductive cells (spermatozoa and eggs). The gametes are haploid and contain a single copy of each chromosome.
Gene: the fundamental physical and functional unit of inheritance. It is a DNA segment made of a transcribed region and a regulatory sequence. Typically it contains the message for protein production; there are nevertheless some genes producing mRNAs that will never be translated into proteins.
Genetic code: the code in which three DNA bases correspond to an amino acid. The bases transcripted from DNA to mRNA are read three by three. The reading of these triplets determines the amino acid sequence of a protein, that will be synthesized according to the genetic information contained in the mRNA.
Genetic collection: collection of DNA clones starting from a donor DNA and representative of the donor's entire genome.
Genetic marker: identifies an allele associating with a character under investiagtion (e.g. a particular disease) and can therefore be used in diagnostics.
Genetics: the study of the hereditary component of living beings' characters. It is the process allowing the transmission of specific characters from parents to children through several generations.
Gene therapy: an experimental procedure aimed at substituting, manipulating or supporting genes that do not work or work badly with genes which work correctly.

Genome: all the genetic material (DNA) of an organism. The human genome is made up of 3 billion base pairs, including about 25,000 expressed genes and a majority of non-coding sequences.
Genomics: the study of the genome and its products (RNA and proteins). Genomics uses some laboratory techniques aimed at comprehending how DNA information is converted into the biological processes of an organism.
Genotype: the individual's genetic organization in respect to a specific gene or group of genes or the whole genome. The genotype remains unchanged (with some exceptions) for the individual's entire life.
Germinal cells: cells responsible for gamete production; in human beings they are the egg and spermatozoon. The germinal cells are haploid and contain 23 chromosomes.
Germinal mutation: alteration occurring in germinal cells and/or cells derived from them, and transmitted to the next generation.
Guanine (G): a nitrogen purine nucleobase contained in both DNA and RNA. In DNA it pairs with cytosine.
Haploid: cell containing half the number of chromosomes characteristic of the species. In humans the egg and spermatozoon contain 23 chromosomes.
Haplotype: a set of single alleles on a single chromatid that are statistically associated, since they are transmitted as a single block through the genealogical tree. SNP haplotypes are investigated by the human HapMap project. The high number of possible alleles for each genic locus allows every single individual to be characterized by the combination of alleles in the different loci.

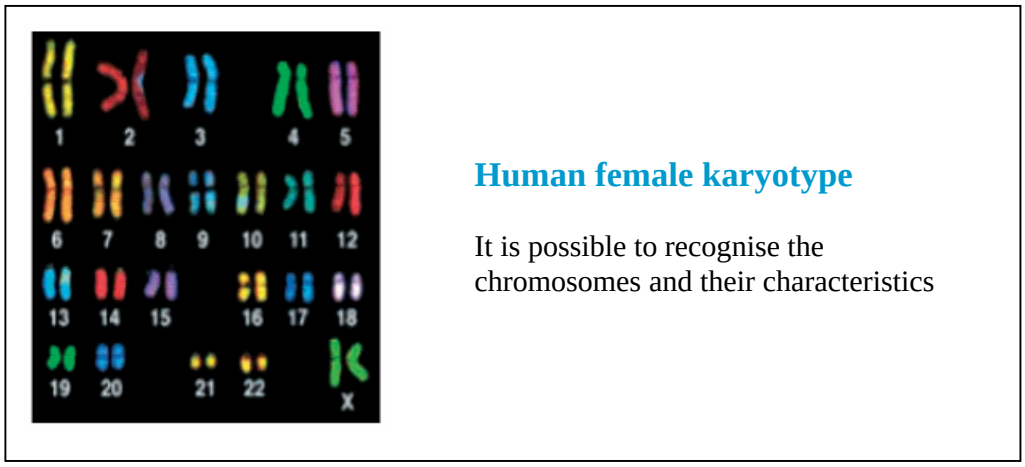


Hereditariness: phenomenon in which certain features are transmitted through generations.
Heterozygote: individual carrying two different alleles of a specific gene on the two chromosome copies.
Homologue chromosomes: two copies of the same chromosome within a cell, one inherited from the mother and the other from the father.
Homozygote: individual carrying two identical alleles of a specific gene on the two homologue chromosomes.
Hot-spot: peculiar DNA sequence with a high frequency of recombination (crossing-over) or mutation.
Hybrid: heterozygous individual, generated by crossbreeding from two parents with different genotypes.

Hybridization: the use of a radioactive or in any case marked probe, such as a single DNA strand, to localize a sequence, gene or mRNA within a cell or tissue. Localization is made possible by the binding between the probe and the gene or RNA, because of complementary sequences.

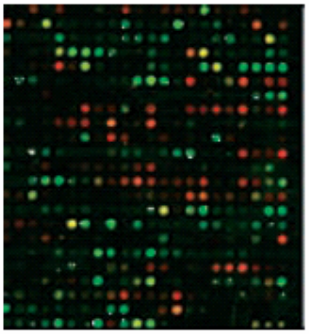


Intron: a non-coding DNA sequence within a gene, removed during mRNA maturation (see **splicing**) and thus not translated into a protein.
Karyotype: the entire range of features identifying a chromosome set. In particular: number of chromosomes, their relative size, length of the chromosome arms, position of the centromere and other features.



Linkage: the tendency of two genes to be inherited together because of their physical proximity on a chromosome. Measured as frequency of recombination: the closer two genes are on a chromosome, the lesser the possibility that they will separate during crossing-over. Linkage analysis uses this feature to identify, by means of specific DNA sequences (markers), disease genes transmitted within some large families.
Locus: unique chromosome localization (literally a “place”), defining the position of a single gene or a certain DNA sequence.
Mapping: physical localization of genes related to a disease on particular chromosomal regions. Consists in defining on which chromosome and in which part of it the gene lies. This is the necessary prerequisite for complete gene identification and cloning.
Megabase (Mb): unit of measurement for the genome, corresponding to one million base pairs.

Meiosis: a process consisting of two successive nuclear divisions, leading to formation of gametes (egg and spermatozoon) and a decreased number of chromosomes, from two copies to just one in each cell.
Microarray: technology for analyzing thousands of genes on a chip, to study their expression and function at the same time (see also **microchip**).
Microchip: small silicon support, on to which fluorescent single-stranded DNA sequences (probes) are spotted. When the probe meets a complementary sequence in the biological sample, it binds to it and emits a signal (microarray hybridization). This allows a simultaneous analysis of several genes and their expression in specific tissues.



Fluorescence signals from a chip

The colour and intensity of the signal are proportional to gene expression

Mitosis: cellular and nuclear division giving birth to two cells identical to the parental one.
Modifying gene: a gene modifying the phenotypic expression (e.g. clinical signs of a disease) of another gene.
Multifactorial diseases: diseases originating from the interaction of one or more genes with environmental factors. Each gene gives its minimal contribution to the expression of the disease phenotype. The most common diseases in Western society, such as tumours, diabetes, asthma, osteoporosis and cardiovascular or psychiatric diseases, belong to this category.
Mutagen: chemical or physical agent which can increase the mutation rate.
Mutation: alteration of the DNA nucleotide sequence. The simplest mutations involve a single nucleotide and are also called "single point mutations".
Northern blotting: a technique that makes it possible to study whether a specific gene is expressed in a tissue and in what quantity. The different mRNAs in the analyzed tissue are extracted and separated by electrophoresis, according to their size, in a denaturing gel (e.g. agarose-formaldehyde). Once separated, RNA molecules are transferred on to a nitrocellulose filter and hybridized with a marked gene probe.
Nucleic acids: the genetic material contained within cells, that is DNA (deoxyribonucleic acid) and RNA (ribonucleic acid).
Nucleotide probe: a single-stranded DNA or RNA sequence, marked with radioactive isotopes or fluorescent chemical dyes, that can be used to identify complementary sequences in genes or RNA. This process is called hybridization (see also **hybridization**).
Nucleotides: the fundamental components of nucleic acids, constituted by a nitrogen base (adenine, guanine, cytosine, thymine, uracyl), a 5-carbon sugar (ribose or deoxyribose) and a phosphate group.
Oligonucleotide: short sequence of synthetic DNA that can be used as a probe (see also **nucleotide probe**).
PCR (Polymerase Chain Reaction): a technique that makes it possible to copy (amplify) specific sequences within a DNA sample and produce a large number of copies (up to 100,000 copies), using the DNA polymerase enzyme. The most common uses are in the diagnosis of genetic

diseases, in producing large quantities of any gene, in pathology and in the study of infectius diseases.

Penetrance: indicates the frequency with which a given genotype will actually result in the corresponding phenotype.

Pharmacogenetics: the study of variability in drug response due to genetic factors. Through the information obtained by genetic studies, it is possible to identify persons responding or not responding to drugs or susceptible to side effects, in order to personalize drug therapies.

Pharmacogenomics: the study of the entire genome and its products (RNA and proteins) to discover and develop new drugs. Discovering the role of genes and proteins in a certain disease permits the design of drugs to overcome it.

Phenotype: the observable features in an individual (e.g. eye colour, hair colour, presence of a disease). The phenotype can be the expression of a certain genotype (genetic structure) or can be influenced by both genetic and environmental factors.

Plasmid: extrachromosomal DNA molecule; it is circular and able to replicate autonomously in bacterial cells; can be used as a vector (see also **vector**) for the diffusion of recombinant DNA molecules.

Polymorphism: the existence, in the population, of many different gene alleles with a significant frequency (more than 1%).

Polyploid: cell or organism with three or more chromosome sets.

Proband: an affected person through whom the family comes to the physician's attention (for example, the individual initially subjected to analysis for a genetic disease, before the rest of the family).

Promoter: the initial portion of a gene where RNA polymerase binds to start transcription of a DNA sequence into mRNA and then into a protein. Greater or lesser promoter activity determines the amount of protein produced (see also **exon**).

Protein: organic compound consisting of amino acids, bound in a specific sequence.

Proteomics: the study of the entire products of the genome, the proteins.

Purine: nitrogen compound made of two ring structures present in DNA and RNA: adenine and guanine.

Pyrimidine: nitrogen compound with a ring structure which is present in DNA and RNA: cytosine, thymine and uracyl.

Recessive: a character that is expressed at phenotypic level only if there are two copies of the related allele in the individual's genotype (the individual is homozygous).

Renaturation: spontaneous pairing between two single-stranded DNAs that re-form the previously denatured double helix.

Restriction enzyme: enzyme cutting the DNA molecule at specific points also known as restriction sites. A technique often used in genetic engineering.

RFLP (Restriction Fragment Length Polymorphism): acronym indicating polymorphisms of the length of restriction fragments. RFLPs are often used as markers (see also **genetic marker**).

Ribosome: small intracytoplasmatic spherical structure where protein synthesis take place. Consists of two subunits made up of rRNA and proteins.

RNA (RiboNucleic Acid): acronym indicating RiboNucleic Acid. A nucleotide polymer made up of 4 nucleobases (adenine, uracyl, cytosine, guanine), a 5-carbon sugar (ribose) and phosphate groups. Synthesized by enzymes which use DNA sequences as templates. RNA is involved in protein synthesis and, according to its function, can be distinguished as:

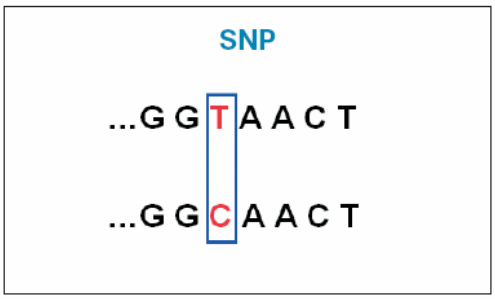
- **mRNA (messenger RNA):** filamentous molecule which specifies the aminoacid sequence of a protein.
- **rRNA (ribosomal RNA):** molecule with a globular conformation representing up to 65% of ribosome weight and participating in protein synthesis.
- **tRNA (transfer RNA**): small RNAs with a globular structure, each of them transporting a specific aminoacid to the ribosome. In protein synthesis they act as adapters, thanks to the

anti-codon, between the mRNA triplet (codon) and an aminoacid. In bacteria there are 31 different tRNA types, in humans there are 48 – fewer than the 61 triplets of the genetic code. A single anti-codon can often bind "synonymous" triplets (coding for the same amino acid), differing only in their third base – e.g. GGG,GGC,GGT (glycine).

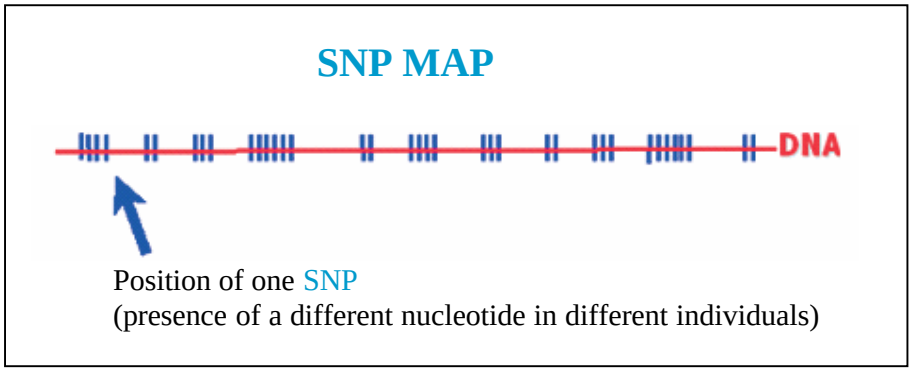
Sequencing: the method to identify the nucleotide sequence of a purified DNA fragment.

Simple diseases: also known as Mendelian diseases, these are diseases caused by a mutation (never the same) in a single gene. Thousands of rare diseases such as cystic fibrosis, Duchenne muscular dystrophy and thalassemia belong to this category. (See also **multifactorial diseases**).

SNP (Single Nucleotide Polymorphism): variation of a DNA single nucleotide base. In humans about 9 million SNPs have been identified. SNPs generally do not cause a disease, but determine variability among individuals.



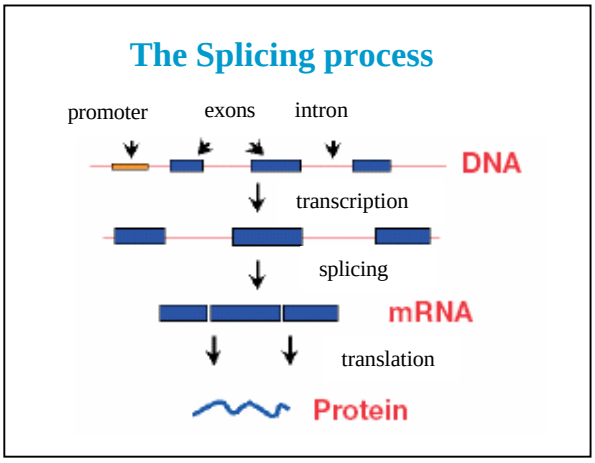
SNP map: the localization of single nucleotide polymorphisms along the DNA (see **SNPs** and **Haplotype**).



Somatic mutation: alteration occurring in a non-germinal (somatic) cell and thus not transmitted to the next generations. Many somatic mutations occur after birth because of environmental factors (e.g. the majority of mutations causing tumours).

Southern Blotting: a technique that makes it possible to isolate from the entire genome DNA fragments recognized by a specific probe after cutting with appropriate restriction enzymes.

Splicing: the process forming mRNA, eliminating the introns and re-binding the exons (see also **alternative splicing**).



Stem cells: primitive cells, undifferentiated, that are able to reproduce and differentiate. These cells are typical of embryonic tissues, but are now known to be present in every tissue of an adult organism, albeit in very limited numbers.
Telomere: the terminal portion of chromosomes. Made up of a series of extremely well conserved short repetitions, occurring one after the other (in tandem), added by the telomerase enzyme to maintain the structural integrity of the chromosome and ensure replication of its final segments.
Teratogen: agent interfering with the normal processes of embryonic development.
Thymine (T): a pyrimidine nitrogen nucleobase contained in the DNA. Pairs with an adenine.
Transcription: process aimed at copying the information carried by a DNA sequence in a complementary RNA sequence by means of the RNA polymerase enzyme (see also **splicing**).
Transgenic: animal or plant organism whose genome has been modified through recombinant DNA techniques.
Translation: process aimed at translating the information carried by mRNA into the amino acid sequence for protein synthesis. This occurs in the ribosome.
Triplet: three base pairs forming a codon.
Twins, dizygote or heterozygote: dizygote twins originate when two different spermatozoa fertilize two oocytes, and can differ in sex. They are simply siblings conceived simultaneously and 50% identical from a genetic point of view.
Twins, monozygote or identical: monozygote twins derive from one egg fertilized by the same spermatozoon. They are thus genetically identical.
Uracyl (U): a pyrimidine nitrogen nucleobase contained in the RNA.
Vector: plasmid DNA, bacteriophage (bacterial virus) or animal virus, used in cloning techniques to spread the DNA (or cDNA) concerned into bacterial or animal cells.
Western blotting: a technique involving electrophoretic separation of proteins, their transfer into a filter and incubation with a marked antibody, capable of revealing a specific protein.
Zygote: cell originating from the fusion between an oocyte and a spermatozoon. Is diploid and will divide by mitosis until it gives rise to a differentiated organism.

Bibliography

1. Erice Declaration on ethical principles of pharmacogenetic research
   http://www.genomica.net/ETICA/dichiarazione_erice.htm

2. Guidelines proposal for the evaluation of a pharmacogenetic experimentation
   http://www.sigu.net/e107_files/downloads/lineeguida/LGFG.pdf

3. Recommendation of the Committee of Ministers to member states on research on biological material of human origin
   http://www.coe.int/t/e/legal_affairs/legal_co-operation/bioethics/texts_and_documents/Rec_2006_4.pdf

4. Guidelines for the institution and the accreditation of biobanks,
   National Committee of Biosafety and Biotechnologies
   http://www.governo.it/biotecnologie/documenti/7.biobanche.pdf

5. From pharmacogenetics to pharmacogenomics, National Committee for Bioethics
   http://www.governo.it/bioetica/pareri.html

6. Document of the teamwork on populations genetic census instituted by the National Committee on Biosafety and Biotechnologies
   http://www.governo.it/biotecnologie/documenti/6.screening.pdf

7. Concept paper on biobanks: pharmacogenetics and pharmacogenomics, EMEA
   http://www.emea.eu.int/pdfs/human/pharmacogenetics/680605en.pdf

8. Our inheritance, our future- realising the power of genetics in the National Health System. Department of Health NHS UK
   http://www.dh.gov.uk/PolicyAndGuidance/HealthAndSocialCareTopics/Genetics/fs/en

9. Agreement among the Ministry of Health, the regions and the Autonomous Provinces on the document: “Guidelines for medical genetics activities”. Permanent Conference for the relationships among the State, the Regions, the Autonomous Provinces of Trento and Bolzano. Official Gazette 224, 23rd September 2004.
   http://pbi.wp.ebixtrade.it/site/pbi_wp_ebixtrade_it/Normativageneticamedica.pdf

10. Draft guidelines for quality assurance in molecular genetic testing, OECD
    http://www.oecd.org/dataoecd/43/26/37103271.pdf

11. Position paper on terminology in pharmacogenetics, EMEA
    http://emea.eu.int/pdfs/human/press/pp/307001en.pdf

12. 25 Recommendations on the ethical, legal and social implications of genetic testing; European Committee
    http://europa.eu.int/comm/research/conferences/2004/genetic/pdf/recommendations_en.pdf

13. European Directive 98/44 on the patentability of biotechnology inventions
    http://www.europarl.eu.int/comparl/tempcom/genetics/links/directive_44_en.pdf

14. Online Mendelian Inheritance In Man
http://www.ncbi.nlm.nih.gov/Omim/mimstats.html

15. Pharmacogenetics Ethical Issues; Nuffield Council on Bioethics
http://www.nuffielbioethics.org

16. Pharmacogenetics-towards improving treatment with medicines. CIOMS,
www.cioms.ch/frame_pharmacogenetics_febr_2005

17. Genetic exceptionalism and legislative pragmatism. Rothstein MA, Hastings Centre Report, July-August 2005
www.thehastingscenter.org

18. Pharmacogenetics and future drug development and delivery.
Roses A., The Lancet 2000, vol. 355, page 1358-61

19. Legislative Decree 30th July 2003, n.196 "Code concerning the protection of personal data"
http://www.garanteprivacy.it/garante/navig/jsp/index.jsp?folderpath=-Normativa%2FItaliana%2FI+Codice+in+materia+di+protezione+dei+dati+personali

20. Authorization n. 2/2005 for the treatment of data revealing the health state and the sexual life
http://garanteprivacy.it/garante/doc.jsp?ID=1203946

21. General measure - 31st March 2004 – concerning situations to be excluded from the duty of notification
http://www.garanteprivacy.it/garante/doc.jsp?ID=1203946

22. Haplotype map of the Human Genome. The international HapMap consortium, Nature 2005, vol. 437, page 1299-1320

23. Searching for genetic determinants in new millennium. Risch NJ, Nature 2000, vol. 405, page 1299-1320

24. Genome-wide association studies: theoretical and practical concerns. Wang WJS et al., Nature Reviews Genetics 2005, vol. 6, page 109-118.

25. Efficiency and power in genetic association studies. De Bakker P.I.W et al., Nature Genetics 2005, vol. 37, page 1217-1223.

26. Genetic Biobanks-guidelines, Italian Society of Human Genetics and Telethon
http://sigu.accmed.org/sigu/html/documenti/proposta_linee_guida_biobanche_genetiche _2003.pdf

27. Data storage and DNA banking for biomedical research: technical, social and ethical issues
European Society of Human Genetics EJHG (2003) 11, supp 2, S8-S10
http://www.eshg.org
http://www.nature.com/ejhg/journal/v11/n2s/pdf/5201115a.pdf

28. Public Population Project in Genomics (P3G)
http://www.p3gconsortium.org

29. http://www.ncbi.nlm.nih.gov/SNPS; http://snp.cshl.org/

30. http://www.hapmap.org/thehapmap.html.en

31. Enforcement of European Directive on the patentability of biotechnological inventions in Italy: Law of 22nd February 2006, n.78.
http://www.filodiritto.com/index.php?azione=visualizza&iddoc=167

32. The ethics of patenting DNA. Nuffiels Council on Bioethics
http://www.nuffieldbioethics.org/go/publications/latest_30.html .

33. Research involving Human biological materials: ethical issues and policy guidance; US National Bioethics Advisory Commission (NBAC)
http://www.georgetown.edu/research/nrcbl/nbac/hbm.pdf