Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

18-May-2005

English - Or. English

DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INDUSTRY
COMMITTEE FOR SCIENTIFIC AND TECHNOLOGICAL POLICY

DSTI/EAS/STP/NESTI(2005)8/FINAL
Unclassified

Working Party of National Experts on Science and Technology Indicators

A FRAMEWORK FOR BIOTECHNOLOGY STATISTICS

JT00184341

English - Or. English

FOREWORD

In December 2004, this report was presented to the Working Party of National Experts on Science and Technology Indicators (NESTI), the Working Party on Biotechnology and experts of the *Ad Hoc* Meetings on Biotechnology Statistics as part of their work on biotechnology statistics. It was recommended to be made public by the Committee for Scientific and Technological Policy in May 2005.

The report was prepared by the OECD Secretariat. It is published under the responsibility of the Secretary-General of the OECD.

TABLE OF CONTENTS

CHAPTER 1: AIM AND SCOPE OF THIS FRAMEWORK

The development and application of biotechnology has the potential for far-reaching economic, social and environmental impacts. It is therefore important to outline a statistical framework to guide the measurement of biotechnology activity. This Framework is intended to provide the basis for statistical compilation work within OECD member countries and those non-member countries wishing to adopt the standards.

The focus of the Framework is on biotechnology R&D and the application of biotechnology techniques to produce goods or services. For simplicity, these are referred to in this Framework as *key* biotechnology activities. End uses of biotechnology, for instance the use of products produced using biotechnology in manufacturing, agriculture or private consumption, are of increasing policy interest but are beyond the scope of this document. However, many of the statistical standards articulated here will be relevant to such uses. The Biotechnology Statistics Conceptual Model in Chapter 2 illustrates the distinction between key activities and end uses.

Under the auspices of the OECD's National Experts on Science and Technology Indicators (NESTI) group, five *Ad hoc* Meetings on Biotechnology Statistics have been held to date. The Biotechnology Statistics Framework is based on the methodological work produced by these meetings (held from 2000 to 2004). It is hoped that publication of the Framework will encourage further statistical work in this field and stimulate debate leading to further improvements in biotechnology statistics.

The Biotechnology Statistics Framework includes the following components:

- Information on concepts, units and definitions for statistical purposes.
- An articulation of user needs and how these relate to the statistical material in the Framework.
- Guidelines for data collection, including a model question on biotechnology R&D, a model survey of key biotechnology activities, and related methodological information.
- Classifications to support the definitions and/or describe statistical units and data.
- Links to other relevant manuals.
- A glossary of terms.

CHAPTER 2: BASIC CONCEPTS AND DEFINITIONS

Introduction

Biotechnology encompasses several different research technologies or methods and several sectors or fields of application. As an example of multiple applications, recombinant DNA technology can be used to produce large molecule medicines in the pharmaceutical sector, create new crop varieties in the agricultural sector, or create micro-organisms that produce industrial enzymes for the chemical sector. The variety of methods plus the range of applications can lead to large differences in how survey respondents might interpret questions on "biotechnology". To avoid this problem, biotechnology must be carefully defined in order to produce reliable and comparable statistics and indicators.

The focus of this framework is on technologies and methods for biotechnology, although the model questionnaire (Annex 2) includes a question on different biotechnology applications. The Framework defines research technologies and methods as biotechnology *techniques* which may lead to a range of biotechnology *applications*.

The statistical definitions of biotechnology techniques were developed through extensive consultation between the OECD secretariat, participants at the *Ad hoc* Meeting on Biotechnology Statistics, the Working Party on Biotechnology (WPB), the National Experts on Science and Technology Indicators (NESTI), and several other expert groups. The Framework includes both a single definition of biotechnology and a list-based definition of biotechnology techniques. Both are necessary to obtain reliable measures of biotechnology activities. Full definitions of each biotechnology technique are given after an overview of the conceptual model for biotechnology statistics.

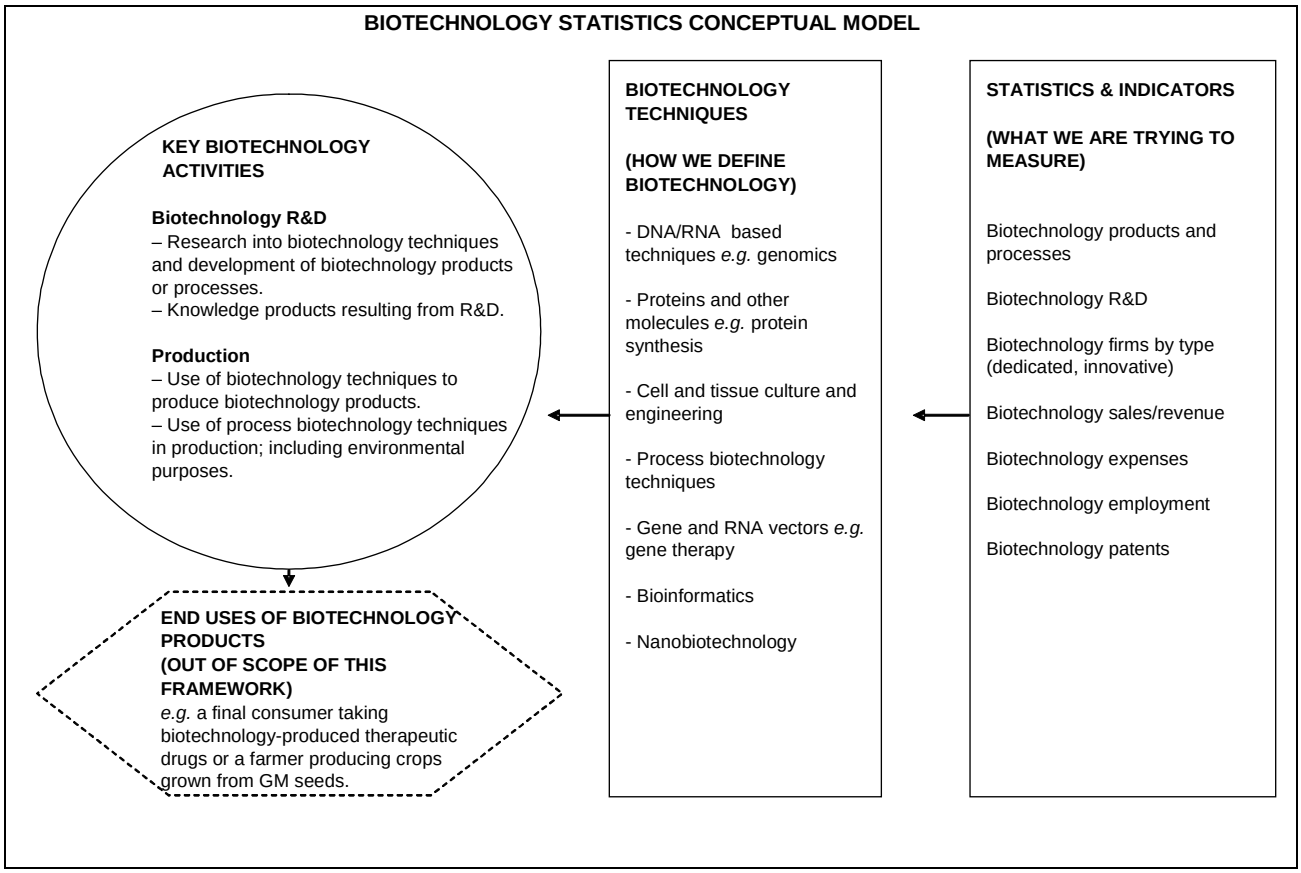
A conceptual model for biotechnology measurement

There are several areas of interest for the collection of statistics on biotechnology techniques that reflect the use of these techniques by firms. Biotechnology techniques can be:

- Studied (basic or applied research) and developed (experimental development).
- Commercialised (*e.g.* the knowledge can be sold or acquired through licensing different forms of intellectual property).
- Used for product or process development.
- Used in production of biotechnology goods and services.
- Resulting biotechnology products can be used by both firms and individual consumers.

The impacts of these activities can be economic in nature (*e.g.* reduction in business costs or improvements in product and process characteristics), social (*e.g.* health improvements) or environmental (*e.g.* reduced biodiversity or more environmentally-friendly manufacturing processes).

The following diagram provides a conceptual model for biotechnology statistics. Note that it lies within a broader conceptual model covering science and innovation statistics generally. The circle in the top left-hand corner includes the key activities that are the focus of this Framework: biotechnology R&D and the use of biotechnology techniques to produce goods or services. These activities produce end products (in the dotted hexagon) that are outside the current scope of this Framework. End users, either firms or individuals, purchase biotechnology products and use them without further modification, although a firm may use biotechnology products as an input for manufacturing, agriculture or energy production. The central rectangle includes seven main biotechnology techniques that are the focus of biotechnology R&D or which are used in production. The right-hand rectangle includes the different aspects of biotechnology that the Framework is designed to measure. These include biotechnology techniques used in the key biotechnology activities.



It is important to note that some firms c

an be engaged in key biotechnology activities and also use biotechnology products. In such cases, only the firm's key biotechnology activities are covered in this Framework. It is also possible for end uses of biotechnology products to feed back to key activities, for instance GM crops can be used as inputs to manufacturing processes based on biotechnology.

The boundary between key activities and some end uses will not always be clear. This is likely to be especially true for some biotechnology processes with applications across a range of industries (for instance, environmental remediation or pollution control activities). In such cases, the activity is an end use if it requires the simple purchase of biotechnology products (which could be either goods or services). For instance, if a mining company employs a contractor to clean up a mine site and the contractor uses

biotechnology products to achieve that, then the mining company would be considered to be an end user. If a municipal authority treats sewage with biotechnology-produced enzymes purchased from a biotechnology firm, then the authority is also an end user. Similarly, the use of biotechnology inputs (such as GM crops) by a manufacturer may require changes to the production process, but the manufacturer is still an end user if the production process does not itself use biotechnology.

Examples of biotechnology techniques, products and processes

Table 1 illustrates the conceptual model through several examples of biotechnology techniques and resulting biotechnology products:

Table 1. Examples of biotechnology techniques, applications, and end uses

Biotechnology technique	Production/application	Product and end use
Develop genetically enhanced or modified microbes or fungi to make enzymes	Produce enzymes, such as proteases, lipases and amylases which remove stains	Enzymes for use as brightening and cleaning agents in detergents[1]
Develop genetically enhanced microbes to make enzymes	Produce enzymes which selectively degrade lignin and break down wood cell walls during pulping	Enzymes for use in paper bleaching[1]
Develop genetically enhanced organisms to produce enzymes	Enzymes that convert crop residues (stems, leaves, straw, and hulls) to sugars that are then converted to ethanol	Ethanol fuel for use in transportation[1]
Use of biomarkers and other biotechnologies to identify genes in wild varieties that confer improved characteristics and their use in conventional breeding programmes	Develop fungal resistance in tomato plant varieties, drought and pest resistance in rice for West African growing conditions	Improved seed varieties for use in agriculture
Use of rDNA technology to transfer genes from one species to another	Develop pest resistant cotton and soybeans containing a gene to produce the *Bacillus thuringiensis* toxin	Improved seed varieties for use in agriculture
Use of rDNA technology to produce large molecule drugs	Produce algucerase rNDA to treat Gaucher's syndrome, human protein C to treat venous thrombosis, etc.	Therapeutic medicines with new modes of action for use by patients
Lipid and pegylation techniques for improved drug delivery	Modify interferon to reduce injection site reactions and frequency of injections	Therapeutic medicines with improved half-lives and reduced side effects for use by patients
Identification and genetic modification of plant genes for tolerating heavy metal contaminants	Develop plant varieties that can absorb soil or water contaminants such as cadmium or zinc	Use of plant varieties in phytoremediation to clean contaminated soils or groundwater

1. Example taken from report: New Biotech Tools for a Cleaner Environment, Biotechnology Industry Organization, 2004.

Defining biotechnology

It is strongly recommended that collection agencies provide survey respondents with both the single definition of biotechnology and the list-based definition.[1] It is further recommended that statistical agencies provide an "Other (please specify)" category when using the list-based definition categories as question items. This will allow respondents to report biotechnology techniques that fit the single but not the list-based definition and will thus assist in updating the list-based definition. An example of such use of an "Other (please specify)" option is shown in the model questionnaire presented in Annex 2.

The single definition

The provisional single definition of biotechnology is deliberately broad. It covers all modern biotechnology but also many traditional or borderline activities. For this reason, the single definition should **always** be accompanied by the list-based definition which operationalises the definition for measurement purposes. The single definition is:

> *The application of science and technology to living organisms, as well as parts, products and models thereof, to alter living or non-living materials for the production of knowledge, goods and services.*

The list-based definition

The following list of biotechnology techniques (see Box 1) functions as an interpretative guideline to the single definition. The list is indicative rather than exhaustive and is expected to change over time as data collection and biotechnology activities evolve.

Box 1. The list-based definition of biotechnology techniques

DNA/RNA: Genomics, pharmacogenomics, gene probes, genetic engineering, DNA/RNA sequencing/synthesis/amplification, gene expression profiling, and use of antisense technology.

Proteins and other molecules: Sequencing/synthesis/engineering of proteins and peptides (including large molecule hormones); improved delivery methods for large molecule drugs; proteomics, protein isolation and purification, signaling, identification of cell receptors.

Cell and tissue culture and engineering: Cell/tissue culture, tissue engineering (including tissue scaffolds and biomedical engineering), cellular fusion, vaccine/immune stimulants, embryo manipulation.

Process biotechnology techniques: Fermentation using bioreactors, bioprocessing, bioleaching, biopulping, biobleaching, biodesulphurisation, bioremediation, biofiltration and phytoremediation.

Gene and RNA vectors: Gene therapy, viral vectors.

Bioinformatics: Construction of databases on genomes, protein sequences; modelling complex biological processes, including systems biology.

Nanobiotechnology: Applies the tools and processes of nano/microfabrication to build devices for studying biosystems and applications in drug delivery, diagnostics etc.

1 . A Statistics Canada study has shown that differences in the results of biotechnology surveys can occur as a result of different interpretations of the meaning of biotechnology (Rose, 2000).

A glossary of terms used in the list-based definition of biotechnology can be found in Annex 1 of the Framework.

Other relevant definitions

In addition to definitions of biotechnology techniques, other definitions are required to cover basic activities, actors, and investments. Thus, this Framework recognises the following terms, with meanings as noted.

Biotechnology product – defined as a good or service, the development of which requires the use of one or more biotechnology techniques per the list-based and single definitions above. It includes knowledge products (technical know-how) generated from biotechnology R&D.

Biotechnology process – defined as a production or other (*e.g.* environmental) process using one or more biotechnology techniques or products.

Biotechnology active firm (enterprise) – defined as a firm engaged in key biotechnology activities such as the application of at least one biotechnology technique (as defined above) to produce goods or services and/or the performance of biotechnology R&D (as defined below).

In the context of the definition of a biotechnology active firm, it should be noted that a firm is a single legal entity and is thus the smallest legal unit for which financial accounts are maintained. This is usually defined as an enterprise. It is not the group of legal units under common ownership, sometimes referred to as an enterprise group in statistical terms; nor is it a single physical location, sometimes referred to as an establishment.

Within the statistical framework, the biotechnology active firm is the statistical unit for which statistical data are compiled. It is not necessarily the reporting unit, which is the entity that provides statistical information. While statistical units will generally be the same as reporting units, reporting units may be either lower level or higher level units within the firm or enterprise group.

Dedicated biotechnology firm – defined as a biotechnology active firm whose **predominant** activity involves the application of biotechnology techniques to produce goods or services and/or the performance of biotechnology R&D.

Innovative biotechnology firm – defined as a biotechnology active firm that applies biotechnology techniques for the purpose of implementing new or significantly improved products or processes (per the *Oslo Manual* (OECD, 1997) for the measurement of innovation). It excludes end users which innovate simply by using biotechnology products as intermediate inputs (for instance, detergent manufacturers which change their formulation to include enzymes produced by other firms via biotechnology techniques).

Biotechnology research and experimental development (R&D) – defined as R&D into biotechnology techniques, biotechnology products or biotechnology processes, in accordance with **both** the biotechnology definitions presented above and the *Frascati Manual* for the measurement of R&D (OECD, 2002).

Biotechnology sales/revenue – defined as the revenue generated from the sale (or transfer) of biotechnology products (including knowledge products) as defined above. It is thus generally a subset of the total revenue earned by biotechnology firms.

Biotechnology expenses – defined as an expense incurred in the generation of biotechnology revenue. It is thus generally a subset of the total expenses incurred by biotechnology firms.

Biotechnology employment – defined as the employment involved in the generation of biotechnology products as defined above. For ease of collection, it is suggested that employment be measured in terms of staff numbers rather than hours worked. However, where countries prefer, they can collect this information in terms of full-time equivalents, consistent with an R&D survey approach (as outlined in the *Frascati Manual*).

Biotechnology patent – defined as a patent belonging to a defined list of International Patent Classification codes (these are defined in Chapter 5).

As this Framework views biotechnology as a set of techniques with applications in many sectors, no attempt has been made to classify biotechnology as a separate aggregation of industries.

CHAPTER 3: BIOTECHNOLOGY STATISTICS FOR POLICY NEEDS

Identifying user needs

Biotechnology is a set of transforming and enabling technologies that attract significant policy interest. There is therefore a need for relevant statistics to facilitate informed policy debate.

Some of the most important policy issues for biotechnology relate to knowledge inputs such as R&D and skilled personnel; the movement of knowledge across firm, regional and national boundaries; finance; incentives; and markets for biotechnology products. Relevant indicators include linkages between the parties involved, the extent of national and international collaboration and overseas investment, adequacy of capital markets and availability of venture capital, availability of skilled personnel and the ability to hire personnel from abroad, protection of intellectual property, incentives for research and development, adequacy of research and production facilities, and the impact of regulatory regimes.

Biotechnology can also be seen as the application of a set of techniques ultimately leading to the production of a diverse range of goods and services. These activities and outputs will have economic, social and environmental impacts. Economic impacts include changes to the general industrial structure of a country's economy and its international competitiveness. Social impacts of biotechnology will be particularly felt in areas like human health. The environmental impacts will influence resource sustainability and biodiversity.

According to Arundel (2003), statistical indicators of relevance to biotechnology policy cover four broad areas as follows:

- "Supporting biotechnology research: The two main types of government programmes to support biotechnology research are either direct funding of research by the public research sector or direct (research grants) and indirect (tax deductions for research expenditures) funding of research by the private sector. Government funding of both public and private biotechnology research can be substantial, with almost half of all biotechnology R&D in the United States funded by Government. The US Federal Government spends an estimated USD 6 billion on biotechnology research (Senker and Zwanenberg, 2000), while Europe, Australia and Canada combined spent approximately USD 3.4 (PPP) billion on biotechnology in 1997 (van Beuzekom, 2001)." Relevant indicators for public funding of biotechnology research consist of both basic data on public R&D spending in biotechnology and intermediate output measures of public biotechnology research, such as patenting by public research institutes and citations to public research papers.
- "Diffusing biotechnology knowledge and expertise: Many public policies provide incentives for collaboration in order to diffuse knowledge and expertise among different actors. These include subsidies to private firms to contract research out to public institutes, passive incentives to increase the number of contacts between public research and private firms, and research subsidies for private firms that require collaborative networks." Relevant indicators include public-sector and other patents, citations, alliances and licensing activities.
- "Commercialising biotechnology research: Policy makers in several OECD countries believe that firms in their country lag behind the United States in their ability to commercialise national

biotechnology research efforts. The result has been the development of a variety of policies to encourage commercialisation. Several EU countries including Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Sweden and the United Kingdom provide subsidies or grants to increase seed and start-up capital for small biotechnology firms, including university spin-offs and start-ups." Relevant indicators include patents and other technical know-how (TKH), venture capital investment, alliances, sales and employment.

- Encourage the application in production and end uses of biotechnology: "Policies to encourage the application and use of biotechnology include procurement, demonstration projects, information programmes, technology adoption subsidies and appropriate regulatory approval systems. Many of these, such as information programmes, are targeted towards small firms." Relevant indicators include sales revenue, particularly by field of application (health, agriculture, industrial processing etc.), types of biotechnology use and exports of biotechnology products.

Statistical indicators in general need to be unbiased and relevant for assessing policy actions and public benefits. This is also true of biotechnology indicators and may be more challenging because of factors such as the changing nature of biotechnology techniques and relatively complex statistical concepts. Various quality aspects of biotechnology statistics are discussed in Chapter 4.

Given the current emerging state of biotechnology, the focus of this Framework is directed towards developing statistics and indicators which focus on key biotechnology activities as described in Chapter 2. The first three of Arundel's four areas are relevant to key biotechnology activities, whilst the fourth is relevant to both key activities and end uses (as described in Chapter 2).

Available indicators

Table 2 (adapted and updated from Arundel 2003) shows biotechnology indicators which are available for 2000 to 2004 for at least one OECD country. A "high" availability rating is used when the indicator is available for 15 or more OECD countries and a "low" rating is used when the indicator is available for three or fewer countries. The table does not list all the variations in the available indicators, but assigns them to major types, such as "patents granted" or "trade in biotechnology/biotechnology exports". The table also shows the main sources for the data. The availability ratings give preference to high quality indicators available from National Statistical Offices (NSOs). For example, an indicator that is available from NSO sources for 10 countries (medium availability) and from private organisations for 16 countries (high availability) is given a medium availability rating if the NSO data are of higher quality than the private data.

Table 2. Biotechnology Compendium indicators: relevance for main policy areas

Indicator	Biotechnology research	Dissemination of knowledge	Commercialisation	Application or use[1]	Impacts[2]	Availability by country[3]	Main data source[4]
Patents granted	✓	✓			?	High	GOV
Patent applications	✓	✓			?	High	GOV
Patent share of worldwide patents[5]						High	GOV
Patent growth rate	✓	✓				High	GOV
Citation share of worldwide citations[5]						High	ACD/PRI
Citation impact		✓				High	ACD/PRI
Venture capital investment	✓		✓			High	NSO/PRI
Total business biotech R&D expenditures	✓					High	NSO
Field trials by trait[6]					✓	High	GOV
Total public biotech R&D expenditures	✓					Medium	NSO
Number of biotechnology firms by field/sector		✓		?		Medium	NSO
Biotechnology alliances and outsourcing		✓	?			Medium	NSO/ACD
Obstacles to commercialisation			✓			Medium	NSO
GM crop area[7]				✓		Medium	GOV
GM crop area by trait[7]				✓	✓	Medium	GOV
Public R&D funding by field	✓					Medium	GOV
Private R&D funding by field	✓					Medium	NSO/PRI/ACD
No. of biotechnology firms by size class			?			Medium	NSO/PRI
Biotechnology revenues/sales			✓	✓	?	Medium	NSO
Biotechnology employees	✓	?	✓		?	Medium	NSO
Types of biotechnology used by firms	✓			✓	✓	Medium	NSO
Funding sources for SMEs	✓		✓			Medium	NSO
Trade in biotechnology/exports			✓	✓	?	Medium	NSO
Technology licensing	?	✓	✓			Medium	NSO
Market approval for health products				✓	✓	Medium	GOV
Biotechnology employees by qualifications	✓	?	?			Low	NSO
Number of public biotechnology institutes	?					Low	GOV
Co-patenting or co-publishing		✓				Low	ACD

Notes: ✓ means that the indicator is of use for the relevant policy area. '?' shows the indicator is of minor value or it could be of use if more details by sector or field were available.

1: Final applications as products or processes, as distinct from the provision of R&D services or equipment.
2: The indicator provides information for assessing productivity improvements or social and environmental benefits.
3: High availability: indicator available for over 15 OECD countries, Medium: available for 4 to 15 countries, Low: available for 3 or fewer countries.
4. GOV = government agencies, NSO = national statistical offices, PRI = private organisations, ACD = academic organisations.
5: Main function of indicator is for comparing national capabilities to other countries.
6: This information should be available for all EU countries, the United States, and Canada.
7: The number of countries for which these indicators are available is limited by restrictions on the use of GM crops.

Indicators and how to collect them

There is not necessarily a one-to-one correspondence between indicators and policy issues. A single indicator can often be used as an input into many policy issues. Furthermore, a particular statistical indicator may not provide all the answers relevant to a particular issue; in many cases several statistical and other types of indicators will be required for analysing particular policy issues.

The first priority is to develop a set of relevant indicators for policies supporting biotechnology development such as through R&D. Many of these indicators can be derived from patents, citations, or private sources that do not require surveys. Information on biotechnology development can come from:

- A complete inventory of firms that perform biotechnology R&D, which can be obtained from either adding one or more questions to R&D surveys or by using multiple alternative sources.
- Data on public and private spending on biotechnology research.
- Indicators on the use of public research subsidies by private firms, such as the percentage of private firms that receive public subsidies or the share of all private sector biotechnology research funded by government.
- Number of employees with biotechnology activities, particularly in large firms where employment data are often unavailable.

If possible, indicators should be collected for specific biotechnology areas, such as health, agriculture, industrial processing, and environmental remediation.

The second priority is based on the assumption that biotechnology is moving from a development to an applications phase. The main priority is for a set of indicators on biotechnology applications by field. This data can only be obtained from surveys, although analyses of GM field test data can indicate the types of GM seeds that should reach the market in two to five years and data on Phase III trials and new drug approvals can provide similar estimates for pharmaceuticals.

The third priority is to develop indicators of the economic, social and environmental impacts of biotechnology. These indicators can play a role in each of the four policy areas discussed earlier. On the simplest level, indicators on benefits can be used to guide policy decisions in R&D investment, commercialisation, procurement, demonstration projects, etc. Some of the most useful indicators on benefits in the health and agricultural sector do not require surveys *e.g.* GM seed sales, analysis of drug reviews by public drug assessment agencies or private organisations (Arundel and Mintzes, 2004). The most difficult area is to develop statistics and indicators for the application of biotechnology to industrial processing. Most available data are based on case studies, but prevalence data (including for potential applications) are required in order to assess the possible benefits of greater public investment in these technologies, and to identify barriers to their adoption. The latter include both general conditions (such as public opposition) that could hamper the adoption of a wide range of biotechnology processes and products, plus barriers that only apply to specific biotechnology products or processes (such as competitive alternative technologies or a lack of seed capital for R&D).

Having established a set of policy issues and indicators that are relevant for biotechnology processes and products, it is then appropriate to consider how those indicators could be collected.

There are both survey and non-survey sources of biotechnology data. The main types of surveys for collecting biotechnology statistics are:

- Standard R&D surveys.
- Standard industry surveys based on a sample of all firms in a sector where biotechnology has potential applications.
- Dedicated surveys of firms that undertake key biotechnology activities.
- Other types of surveys, such as surveys of specific sectors, public research institutes or households.

Each of these survey mechanisms has its own strengths and weaknesses in collecting biotechnology related statistics. Adding questions to existing surveys can be an effective and inexpensive means of providing statistical information on biotechnology. But the information obtained through these means is limited to a few indicators. Dedicated surveys tend to be expensive but allow for more elaborate questions related to biotechnology activities, human resources, the commercialisation process, policy and other impacts, and access to financial capital.

Another important aspect to consider when devising a means to collect statistical indicators is response burden and response rate. Short and simple questionnaires will be cheaper and should achieve higher response rates than longer and more complicated surveys.

In many countries cost or regulatory requirements could prevent NSOs from implementing all of the surveys that would be necessary to obtain the proposed indicators. Fortunately many indicators can be obtained from sources other than surveys.

Table 3 (adapted from Arundel 2003) identifies data sources for a selected sample of biotechnology statistics but does not provide a priority rating for collection activity. The choice of which statistics to collect will depend on both their value and pragmatic considerations, such as the ability of NSOs to add a question to an existing survey, or implement specialised biotechnology surveys. It is hoped that the provision of model questions and a questionnaire in Annex 2 of this Framework will influence NSO activity in biotechnology statistics collection.

Table 3. Examples of sources of biotechnology statistics

Survey type and statistics	Survey questions
R&D Survey	
Biotechnology R&D performers	- Yes or no question on R&D expenditures in biotechnology - Expenditures by biotechnology field
Government performance and funding of biotechnology R&D	- Government expenditures related to biotechnology R&D
Biotechnology Survey - narrow (survey of all firms that are thought to perform biotechnology R&D)	
Effect of patents on research projects	- Questions on impact of patenting and licensing by other firms on the ability to complete and implement research projects
Public subsidies for biotechnology	- Share of firms that receive public subsidies - Share of revenues from public subsidies - Share of research expenditures from public subsidies
Biotechnology Survey - expanded (survey of all firms that are involved in key biotechnology activities for R&D and/or application to produce goods or services)	
Productivity effects from the use of biotechnology	- Basic questions on the effects of process biotechnology on demand for inputs (materials, energy) and labour (skilled, unskilled, etc.)
Human resources for biotechnology*	- Ability to hire biotechnology staff - Barriers to hiring
Commercialisation process*	- Obstacles to commercialisation - Access to financial capital*
Financial characteristics*	- Biotechnology revenues and expenditures, share of total revenue and expenditures due to biotechnology
Trade in biotechnology products	- Value of firm exports of biotechnology products - Share of biotechnology exports out of all exports
Industry Survey (sample of all firms in sectors where key biotechnology activities are likely to occur)	
Application of biotechnology	- Determine firms with biotechnology activities per the single or list-based definition of biotechnology - Collect indicators on end uses of biotechnology products or use in processes
Barriers to the application of biotechnology	- Basic questions on barriers to the general adoption of biotechnology products or processes (including regulations), plus questions relevant to specific types of biotechnology
Other surveys	
Comparing benefits of biotechnology versus competing technology	- Detailed surveys of firms in specific sectors, such as seed firms for agro-biotechnology, industrial processors for environmental biotechnology, etc.
Public biotechnology patenting and licensing: exclusive license share, revenues, spin-offs, etc.	- Survey of public research institutes
Social attitudes to biotechnology	- Household surveys (Eurobarometer, etc.)
GM crop use	- Surveys of farmers on hectares planted with GM crops, plus reasons for non-use of GM varieties
Surveys not necessary or optional	
Therapeutic value of bio-pharmaceuticals	
Sales share of bio-pharmaceuticals	
Investment focus in GM traits	- (Survey of agro-biotechnology firms) - Focus on long-term investment, such as in the planning and laboratory stages
Cost of GM traits in agriculture	- (Survey of agro-biotechnology firms)
Patent data for research trends, collaboration	
Trade in biotechnology goods	

* Some information on these activities can be obtained through narrow biotechnology surveys.

CHAPTER 4: COLLECTION GUIDELINES

Survey data on biotechnology can be obtained by adding questions to existing surveys or by conducting dedicated biotechnology collections. This Framework covers the addition of a question to existing R&D surveys and proposes a model survey for key biotechnology activities. However, it should be noted that some of the following guidelines could be applied to other surveys as well.

Statistics about firms undertaking biotechnology R&D, using R&D surveys

The development of new biotechnology techniques and resulting products or processes generally involves the conduct of research and experimental development. This R&D activity takes place within firms, universities, the government sector and in the private non-profit sector. As most OECD countries have R&D surveys that cover these sectors, it is appropriate to consider how those existing surveys can be modified to enable the measurement of biotechnology R&D. The end result would be the production of internationally comparable information on biotechnology R&D.

At the 2002 National Experts on Science and Technology Indicators meeting, delegates recommended the addition of an optional question to national R&D surveys of the business enterprise sector. Consequently, the sixth edition of the *Frascati Manual* (OECD, 2002) includes an optional biotechnology-related R&D question (for further details, see Annex 4 of the Manual). The *Frascati Manual* presents the internationally recognised methodology for collecting and using R&D statistics. It includes basic concepts, guidelines for collecting data and the classifications to be used in compiling statistics. Please refer to the *Frascati Manual* for information on how to define the business enterprise sector.

Box 2 below presents the recommended question. To improve comparability, it is recommended that countries use the OECD biotechnology definitions (both the single and the list-based definitions presented in Chapter 2 of this document).

Box 2. Recommended biotechnology R&D question

Did the R&D reported above include any biotechnology R&D (see definitions)?	Yes / No
If yes, please provide an estimate of the share of the total intramural R&D expenditure reported earlier that is attributable to biotechnology.	_____%

While the current recommendation covers only the business enterprise sector, for many countries the question could also be applied to other sectors (the government, higher education sector and private non-profit sectors). Those countries are encouraged to use the recommended biotechnology R&D question for applicable sectors and to share their experiences with other participating countries.

Dedicated biotechnology collections – the model survey of biotechnology use and development

By 2000, Canada, France, Japan and New Zealand had developed and implemented dedicated biotechnology use and development surveys (van Beuzekom, 2000). By 2004, the number of countries had

risen to nine, with surveys in Belgium, Canada, France, Germany, Japan, Korea, New Zealand, the United Kingdom and the United States (van Beuzekom, 2004).

The model survey was originally presented to the 2002 Meeting on Biotechnology Statistics following its development by a group of experts from Australia, Belgium, Canada, France, Japan, Spain, and the European Commission. The content has evolved somewhat since then, reflecting more recent national experiences with biotechnology surveys and data.

It is evident that there is considerable overlap between the indicators that this type of survey would produce and indicators produced from innovation surveys, R&D surveys and the more traditional economic activity surveys conducted by many NSOs. Care therefore needs to be taken in the design of such surveys to minimise duplication of statistical collection activity and to ensure that survey frameworks, classifications, standards, questions and results are consistent.

There are two other aspects of such surveys that warrant particular attention. The first relates to the survey scope and, in particular, whether to restrict the scope to manufacturing as per the original recommendation. Most countries aim to include firms undertaking key biotechnology activity as defined in Chapter 2. While a high proportion of such firms will be manufacturers, many will be classified to service industries such as Research and development (ISIC Rev. 3.1 Division 73) and some other service industries such as wholesale, waste management and computer services. Excluded from the survey scope are suppliers of goods and services to biotechnology firms.

The second issue relates to the compatibility of the results from different biotechnology surveys. A Statistics Canada study has shown that differences in the results of biotechnology surveys can occur as a result of different interpretations of the meaning of biotechnology (Rose, 2000). Thus, statisticians will need to be aware of the effect of (possible) different interpretations amongst suppliers of data that can arise from using just the “list-based” definition or just the “single” definition (or another definition altogether). To improve comparability, it is therefore strongly recommended that participating countries use both definitions presented in this Framework – and use them in combination. Where countries wish to add extra biotechnology technique categories to their questionnaire, for international comparability purposes they should aim to produce main aggregates that **exclude** those additional techniques.

It is important to note that country agreement to conduct the model survey does not inhibit the use of additional questions in national biotechnology surveys. The model survey provides a small set of questions that form a basis for compilation of internationally comparable data. Participating countries are therefore encouraged to incorporate those questions, plus the concepts and definitions provided in this Framework, into their national surveys.

Objectives of the model survey

The first objective of the use and development model survey is to enable estimation of the intensity and type of biotechnology activity in firms engaged in key biotechnology activities (as defined in Chapter 2). The proposed model survey uses the provisional statistical definition of biotechnology adopted by the different OECD groups (NESTI, WPB) and measures biotechnology activities per the list-based definition. The high priority indicators identified are: the number and characteristics of biotechnology firms (including by firm size and industry if possible), revenue generated from sales of biotechnology products including technical know-how (TKH), biotechnology R&D expenditures, sources of capital finance, human resources employed, and barriers to biotechnology R&D or commercialisation.

Another indicator that was determined to be a priority, but which was not included in this version of the model survey, deals with information on collaborative arrangements (see the section below on Statistics of interest that are not covered by this Framework).

The proposed questionnaire for the model biotechnology use and development survey is presented as Annex 2 to the Framework. The annex also includes information about the questions and some alternative formulations.

Methodological considerations for a biotechnology use and development survey

Target population

The target of the model survey consists of firms engaged in key biotechnology activities, including R&D and the application of biotechnology techniques. The target population of the survey therefore consists of:

- Manufacturing firms currently undertaking key biotechnology activity (as defined in Chapter 2).

- Biotechnology R&D firms with no product sales and consequently classified by national statistical offices (NSOs) to an R&D service industry category.

- Targeted firms classified to industries other than manufacturing or R&D services (the primary goal is to find firms engaged in key biotechnology activities wherever they are currently classified). These include firms classified to wholesaling, for instance local operations of large foreign pharmaceutical firms, whose local affiliate performs biotechnology research but acts mainly as a wholesale distributor and is therefore classified to the wholesaling sector.

- Some types of services firms are included if they are using biotechnology techniques for the purpose of providing a service. These could include waste management and environmental remediation firms that have developed a process that they then provide to other organisations.

Firms which are to be excluded from the model survey are:

- Services firms that only provide routine contract research (such as diagnostics and testing) or consultancy services.

- Biotechnology equipment suppliers as well as other goods suppliers and firms that only distribute biotechnology products.

- End users of biotechnology products and processes, as described in Chapter 2.

The challenges of constructing population lists

Identification of biotechnology activities for the purpose of conducting surveys is a significant challenge for most countries. In part, the challenge lies in the inherent nature of biotechnology as a diverse set of activities across different sectors of the economy.

Defining exactly which firms are to be included and which excluded requires the development of clear parameters for the populations and an understanding of the nature of biotechnology activities. There are no recognised biotechnology industries in existing industrial classifications and biotechnology activities are

found in many industries. In general, most potential participants in biotechnology surveys are relatively new entities or are firms where biotechnology is a relatively new activity.

Participating countries use a variety of ways of creating population lists for surveys of biotechnology firms. Commonly used are lists of firms known to be engaged in biotechnology activities. Such lists already exist in many countries and are available from sources such as privately-maintained and government directories and lists of beneficiaries of government grants, contracts or assistance programmes. However, there are several possible drawbacks to the use of such lists, including:

- The criteria used to select firms for a list can vary considerably and can be quite different from the criteria used to define biotechnology firms in the model survey.

- The use of lists does not provide any means to assess whether the resulting population of firms engaged in biotechnology is exhaustive or complete.

- The accuracy of the lists may not be good and may be difficult to assess. In addition, the list may not be up to date, which will reduce accuracy.

Methods used in list construction

There are three known key techniques used in the construction of population lists. They are:

- Custom construction of a list using existing information sources.
- Large scale sampling.
- Keyword searches.

The most common method used by participating countries is custom construction of a list using a variety of information sources. Survey practitioners first define the parameters of the intended population and then prepare their lists using some or all of the following sources:

- Biotechnology directories from government and private sources.
- Membership lists of trade organisations.
- Other government departments.
- Academic sources.
- Internet searches.
- Large statistical databases – both private and government.
- Firms responding to previous surveys; these could be previous iterations of a biotechnology survey or other collections, for instance, R&D surveys.
- Published reports.
- Industry experts.
- Government administrative records, including lists of recipients of contracts and subsidies.

The advantage of using existing information sources such as biotechnology directories and association membership lists is that there is generally a good probability that the firms will be involved in biotechnology. However, as discussed above, directories may not be current and their accuracy could be difficult to assess.

Random sampling of biotechnology firms may be used to supplement lists in selected industries. The purpose of this process is to identify potential respondents that could have been missed in the construction of the lists.

The main drawback of large scale sampling is the cost in time, money and response burden, both to the surveyor and respondent. Careful targeting of populations using short questionnaires assists in reducing the time and financial costs of this method. The advantage to this method is that it enables identification of sectors of the economy where biotechnology may be found and an assessment of its intensity in that sector.

Another method used in the construction of lists is a key word search of data bases, including business registers held by NSOs. Firms may be identified as potential respondents based on inclusion of certain key words in their name or description of their activities. Such a key word search could also be used to identify industries that should be targeted for further sampling.

The disadvantage of keyword searches is the potential inclusion of out of scope entities if the key word search is too broad, for example, searching on the word "bio". Further, the name of a firm may not be a good indicator of its activities. Therefore key word searches limited to firm names could miss a large fraction of eligible firms. For example, a key word name search using "bio" and "gen" would only identify 8 (18.6%) of 43 firms that received US marketing approval for a large molecule bio-pharmaceutical up until the end of 2003.

Data quality aspects of biotechnology surveys

The need for biotechnology data users to be able to judge quality and *fitness for use* has become increasingly important with the growing amount of available data and the changing nature of the subject being measured. The OECD (2003) and individual statistical agencies have outlined criteria (or dimensions) for the quality of statistical products.

For the purposes of this Framework, the following quality dimensions are proposed (OECD, 2004*a*):

- Relevance: whether the concept measured corresponds to the concept required.
- Timeliness: the period between the time of data release and the time of the event is vital to users if timely decisions need to be taken.
- Accuracy: the deviation between the target value determined by a perfect process (true parameter) and the value determined by the imperfect process (estimate).
- Accessibility: whether the user can easily make use of the data.
- Comparability: reliable comparison possible over space and time.
- Coherence: whether different sources are based on common definitions, methods, etc.
- Completeness: whether available output reflects all user needs and priorities.

Data quality may be examined via means of a data quality framework that poses a number of questions to be addressed under each of the factors listed above. Full details of such a framework may be found in OECD (2004*a*), however, some of the more important aspects are:

- Scope and coverage. What population is **actually** covered by the data collection and are excluded units different from included units?

- Reporting unit. The reporting unit is important as the information collected will tend to be from the perspective of that unit.

- Classifications, concepts and data items. Are they logical and consistent? Do they match user needs?

- Level of sampling error (applicable to survey samples only). If the confidence interval around an estimate is large, this can affect decisions based on the data. In general, the more disaggregated data become, the greater the level of sampling error. Efficient sample design can significantly reduce sampling error for a given sample size. For biotechnology surveys based on lists, sampling error is likely to be less of a problem than error based on incomplete coverage of biotechnology firms.

- Ability of respondents to provide information. Do survey respondents understand the questions, definitions and concepts that are expressed via survey instruments? Is the requested information likely to be readily available? Is the chosen respondent the best person to provide the information? Could there be recall problems (for instance, for information collected some time after the reference period)?

- Unit and item response rates. The unit response rate is calculated by dividing the number of responding units by the number of units that were selected and were in scope of the data collection. In most instances, an assumption is made that non-respondents would have provided similar information to respondents. However, the characteristics of non-respondents may in fact be quite different to those of respondents, so the data will be biased to reflect those units that responded. Clearly, higher response rates are to be preferred to lower the level of both non-response bias and sampling error. Item response rates refer to individual questions and may be addressed by imputing missing values (see the discussion below on data editing and imputation).

- Data editing and imputation. Editing is the process of checking data records to ensure that they contain valid entries and changing records where they do not, whereas imputation is the process of estimating data for incomplete individual records.

- Timeliness of data. If circumstances are likely to have changed significantly since the reference period, data will not be current. The implications are that data should be collected more frequently and/or released more quickly after collection.

- Coherence. Consistency of classifications, concepts, scope and methodology. If comparisons are to be made (for instance, over time or between countries), major changes in approach to measuring biotechnology can cause problems so should be avoided where possible. Examples include how the level of use of biotechnology techniques is increasing or how expenditure on biotechnology R&D as a proportion of GDP differs between countries.

It is proposed that participating countries provide quality measures as follows:

- Sampling error, where relevant, for key data items, preferably expressed as a relative standard error. The latter is defined as the standard error of the estimate expressed as a proportion (often a percentage) of the estimate.

- Unit response rate expressed as the number of responding units divided by the number that were selected **and** found to be in scope (note that out-of-scope units should be excluded from both the numerator and denominator). The response rate is usually expressed as a percentage.

- Other useful measures include the proportion of the value of key estimates that were estimated by the statistician rather than provided by the respondent. Such estimates could result from unit or item non-response. A simpler but less meaningful measure is the proportion of units for which data items have been altered on the basis of editing or imputation.

Statistics of interest that are not covered by this Framework

Extensions to the model survey

There are a number of questions on highly relevant topics that participating countries have used but which have not been included in this version of the model questionnaire (see Annex 2), either because the questions have only been tested in a limited number of countries or because a large number of detailed questions would be required to collect useful information. Examples include questions on the impacts of biotechnology, the number of development projects within a product pipeline, detailed breakdowns of employment by occupation, staff recruitment problems, collaborative arrangements, and environmental applications.

Regarding collaborative arrangements, country questions are fairly diverse and include collecting information by the type of collaboration and whether those organisations are domestic or foreign. The response categories are either tick-boxes (yes or no) or counts of collaborative arrangements or partners.

Other areas of interest which could be examined in the future include:

- Adding financial information to Question 5 (*Financial characteristics*) to enable the calculation of value-added.

- Collecting further "intensity" details for Question 4 (*Status of biotechnology activities*), for instance expenditure on R&D or revenues associated with each application.

Statistics about end users of biotechnology products or processes

The model survey focuses on the key biotechnology activities of R&D and the application of biotechnology techniques. It is likely that many of the major impacts from biotechnology will come through use of biotechnology products by other firms or individual consumers. While such end uses are beyond the current scope of this Framework, they are acknowledged as having potentially significant economic, social and environmental impacts. Some relevant indicators can be obtained from non-survey methods (see Table 3). Future surveys could include measures of the end uses of biotechnology products and processes. This would require survey development work as well as the development of relevant definitions and classifications.

Statistics about public R&D funding

Recommendations for collecting statistics about public R&D funding are also beyond the scope of this version of the Framework. However, they are seen as highly relevant to policy decisions and could represent a future extension of statistical standards development.

Extension to other sectors

The recommended model R&D question (see Box 2 above) covers only the business enterprise sector. In many countries, significant biotechnology R&D occurs in education and government institutions and it would be of interest to ascertain whether the current model question is suitable for those sectors. If not, then development of relevant questions could be the subject of future statistical work.

The same argument applies to the model survey of biotechnology use and development, some of whose questions could apply to non-business sectors. They include biotechnology R&D details, biotechnology collaborative arrangements and intellectual property protection.

Statistics about social issues

Social issues associated with biotechnology are also of interest. Measures include indicators of peoples' perceptions of impacts of biotechnology on health or the environment. To measure perceptions generally requires a household survey. NSOs typically have a programme of such surveys and it is often possible to incorporate a small number of specific questions into existing surveys. In other cases, countries have developed generalised social or omnibus surveys, which could be utilised for biotechnology data collection.

In some countries, policy agencies associated with biotechnology have conducted their own household surveys (generally via private sector data collectors) to gather indicators about perceptions of specific biotechnology matters. Given the sensitive nature of many of the issues and the need to gather data on opinions very quickly, this strategy will often be the preferred approach for gathering this type of information.

Because national interests and capabilities differ, conducting internationally comparable social surveys may be difficult. Therefore it is left to individual NSOs to decide whether they will conduct surveys or leave evaluation of social attitudes to the private sector.

CHAPTER 5: CLASSIFICATION SCHEMES

Use of classification schemes for biotechnology

Box 1 of Chapter 2 provides a classification scheme for several types of biotechnology techniques. This Chapter provides other classification systems for characterising and identifying biotechnology firms (or institutions), biotechnology activities, and biotechnology applications. These include an industrial classification of the sectors where biotechnology is most likely to be used or researched, the types of institutions that are involved in key biotechnology activities, a size classification for biotechnology firms, patenting classes for biotechnology, commodity classifications, and a system for classifying biotechnology applications. Each classification system is described below.

Industrial classifications

Industrial classifications are used in many statistical collections to determine the scope of the survey and to classify survey results. The International Standard Industrial Classification (ISIC Revision 3.1) has three criteria to determine whether particular activities should be combined or not. These are:

- The character of the goods and services produced.
- The use of those goods and services.
- The inputs, process and technology of production.

Biotechnology, as discussed earlier, is a collection of techniques and hence, in theory, might be recognised by the three aforementioned criteria. However, experience has shown that biotechnology cannot be recognised using existing industry classifications and methods. A future development in this area is the possible inclusion of biotechnology as a subdivision of ISIC class 67 (Scientific research and development) in ISIC Revision 4, scheduled to be released in 2007.

For surveys measuring biotechnology R&D, it may be possible to use the industrial classification in a limited way to restrict the survey scope. While a great deal of biotechnology R&D will be carried out in ISIC class 73 (Research and development), there are likely to be other classes to which biotechnology R&D performers are classified. For countries that use their existing R&D surveys to measure biotechnology R&D data, it is unlikely that the industrial classification will be of any additional help in restricting the scope of the survey.

Pending further analysis of member country data from existing biotechnology surveys, an industry classification is only recommended for adding a biotechnology question to an existing R&D survey or for identifying a sample frame for specialised biotechnology R&D surveys. The industry classification recommended is based on that shown in the *Frascati Manual* 2002 (p. 57 Table 3.1) for the presentation of R&D statistics (see Table 4 below). For many industries, biotechnology activity is likely to be at a low level and therefore the recommended classification has been adapted by excluding two-digit industries that are unlikely to have much biotechnology R&D activity (for example ISIC 22 for publishing and ISIC 34 for automobiles) and expanding others which are of particular interest to the three- or four-digit level.

Countries that collect biotechnology R&D statistics independently of their normal R&D survey may choose to focus on the selected detailed industries included in Table 4. Countries wishing to show an all-industry total could also survey sectors at a higher level of aggregation (*e.g.* Manufacturing, Construction and Services).

Table 4. Proposed industrial classification for use with biotechnology R&D statistics

	ISIC Rev. 3.1 Division/Group/Class
AGRICULTURE, HUNTING, FORESTRY AND FISHING	**01, 02, 05**
MINING AND QUARRYING	**10-14**
MANUFACTURING	**15-37**
Food products and beverages	15
Textiles, wearing apparel, fur and leather etc.	17-19
Paper and paper products	21
Coke, refined petroleum products and nuclear fuel (less refined petroleum products)	23 (less 232)
Refined petroleum products	232
Chemicals and chemical products (less pharmaceuticals)	24 (less 2423)
Pharmaceuticals, medicinal chemicals and botanical products	2423
Non-metallic mineral products, basic metals	26-27
Medical, precision and optical instruments, watches and clocks (less medical and surgical equipment etc.)	33 (less 3311)
Medical and surgical equipment and orthopaedic appliances	3311
Recycling	37
ELECTRICITY, GAS AND WATER SUPPLY	**40-41**
Collection, purification and distribution of water	41
CONSTRUCTION AND SERVICES	**45, 50-99**
Wholesale of agricultural raw materials, live animals, food, beverages and tobacco	512
Computer and related activities	72
Research and development	73
Technical testing and analysis	7422
Human health activities	851
Veterinary activities	852
Sewage and refuse disposal, sanitation and similar activities	90
GRAND TOTAL (all industries)	**01-99**

Institutional sector and type of institution classifications (for biotechnology R&D statistics)

The *Frascati Manual* recognises the existence of the following sectors within an institutional sector classification:

- Business enterprise sector.
- Government sector.
- Private non-Profit sector.
- Higher education sector.
- Abroad.

The *Frascati Manual* recognises the importance of classifying R&D statistics by the type of institution performing R&D. As a key use of biotechnology R&D statistics will be to make comparisons with other types of R&D, it is desirable to use the same classification in this Framework as well. The proposed classification of different types of institutions is as follows:

- Private enterprises:
  - Enterprise not belonging to any group.
  - Enterprise belonging to a national group.
  - Enterprise belonging to a foreign multinational group.
- Public enterprises:
  - Enterprise not belonging to any group.
  - Enterprise belonging to a national group.
- Other research and co-operative institutes.

Public enterprises are distinguished from private enterprises on the basis of control. The SNA 93 makes the following recommendation for the definition of public non-financial corporations:

> "These consist of resident non-financial corporations and quasi-corporations that are subject to control by government units, control over a corporation being defined as the ability to determine general corporate policy by choosing appropriate directors, if necessary. The government may secure control over a corporation:
>
> - by owning more than half the voting shares or otherwise controlling more than half the shareholders' voting power; or
> - as a result of special legislation, decree or regulation which empowers the government to determine corporate policy or to appoint the directors."

A group must be considered as foreign when the main shareholder is a foreign resident with more than 50% ownership and voting power, either directly or indirectly through subsidiaries.

As it will be important to compare biotechnology R&D data with other types of R&D, it is desirable to utilise the same classification list for the measurement of biotechnology R&D. For biotechnology R&D firms surveyed using existing R&D survey vehicles, this differentiation should occur as a matter of course.

Size classification

This classification relates to the size of biotechnology firms and is relevant for both R&D surveys and surveys of key biotechnology activities. The classification of firm size should align with those in the current *Frascati Manual*.

As the number of biotechnology firms in most countries is small, it may be necessary to reduce the number of size categories. This Framework therefore recommends the use of only three categories as shown below:

- 0 – 49 employees (three categories in the *Frascati Manual:* 0, 1-9, 10-49 employees).
- 50 – 249 employees (two categories in the *Frascati Manual*: 50-99, 100-249 employees).

- 250 employees and above (three categories in the *Frascati Manual:* 250-499, 500-999, 1000+ employees).

Where statistically feasible, countries are encouraged to further split the first size category to distinguish the three *Frascati Manual* categories 0, 1-9 and 10-49 employees. This will be more relevant for countries with a high proportion of small and medium-sized enterprises (SMEs) involved in biotechnology activities.

The size classification is also often used to delineate cut-off boundaries in statistical collections. In the case of biotechnology surveys, however, it is proposed that no cut-offs be adopted because of the relatively small number of biotechnology firms and the potential importance of small start-up companies operating in this field.

Patents classifications

Patents can yield information that may not be captured by other indicators. They are important in all technology fields, but probably even more important for new and or specialised fields such as biotechnology. In certain technological areas, such as aerospace, other methods (such as secrecy, lead time, etc.) are used to protect inventions. Whereas in the field of biotechnology, patents are intensively used to protect invention, hence one would expect to capture some of the dynamics of biotechnology activity using patent statistics. Furthermore patents can be important for biotechnology firms as many of them have no activity other than R&D and do not directly exploit their inventions. These firms can use patents to either attract investment or earn income by selling or licensing their patents to other firms.

Indicators based on biotechnology patents may provide some insight into the level of biotechnology activity across countries. However, in order to obtain an accurate picture of such activity using patent information, it is important to develop a robust definition for biotechnology patents. Over the past few years, the OECD has conducted experimental work to develop an operational definition of biotechnology patents using various methodologies.

The goal is to avoid, as much as possible, two types of errors: the inclusion of non-biotechnology patents and the exclusion of relevant biotechnology patents. The definition (described in terms of International Patent Classification (IPC) codes) and given in Table 5, has been reviewed and verified by patent classification experts, whose general conclusion was that it captures a significant proportion of biotechnology patents. Analysis of a sample of patents from Finnish “biotechnology” firms provided further evidence of the robustness of the definition.

Methodology used to identify biotechnology patents

A patent document is a rich source of information, containing specific technical detail, including a list of “claims”, technical classes to which the invention belongs, citations to prior inventions, etc. A large volume of published patent documents is available throughout the world and in recent years around one million patent documents have been published each year. This makes patents one of the largest information sources for tracking innovative activities.

The technical content of the patent document is organised and indexed using patent classification systems, such as the International Patent Classification (IPC) or the US Patent Classification. Patent classification systems are developed and maintained by patent examiners in order to enable them to identify specific topics or technology areas more easily. As part of the examination process, an examiner will assign patent classification codes to the patent specifications (or claims). A patent classification is a hierarchical system divided and subdivided to a detailed level, which theoretically allows the subject of

each patent to be properly classified. In practice, there is some variation due to differences in the preferences of the patent examiners.

The method chosen to identify biotechnology patents was to select a list of IPC codes that encompasses these technologies. The biotechnology IPC codes can be identified using the following methods:

1. Analysis of the IPC classification: this includes scanning the patent classification in a top-down approach, starting at the section level, followed by sub-sections, classes, sub-classes, groups and sub-groups.

2. Keyword search: this includes analysing the patents (normally the titles or the abstract) for a list of keywords associated with the technology. This enables one to identify by statistical analysis the relevant IPC codes wherein these keywords are used most frequently. However key words searches in official documents published by national or regional patent offices are generally less productive, because the legal requirements of disclosure with regard to titles and abstracts are not very strict. If the keyword search is employed to identify the IPC codes for a specific technology, then the search should be performed on a high quality database, such as the World Patent Index.

3. Analysis of patents owned by biotechnology firms: this includes identifying and collecting patent documents of known biotechnology firms in order to perform statistical analysis on the IPC codes allocated by patent examiners.

The first method as outlined above (scanning the patent classification in a top-down approach) was chosen to identify the appropriate IPC codes for biotechnology patents.

The initial phase of the project was carried out by a patent examiner from the Japanese Patent Office in collaboration with the OECD. The result of this work was two alternative definitions of biotechnology.

In order to refine the list of IPC codes for biotechnology patents, a second patent classification expert analysed the whole IPC classification using a top-down approach. This resulted in the modification of the list of IPC codes considered to be in the field of biotechnology. The updated list of IPC codes for biotechnology patents is shown in Table 5. The majority of biotechnology patents are in sub-classes C12M to C12S. For example, these five classes accounted for 82.2% of European Patent Office (EPO) patent applications in 2000 within all classes listed in Table 5, with 55.5% attributable to C12N alone.

A second patent classification expert from the World Intellectual Property Organisation (WIPO) analysed the modified list of IPC codes and concluded that the modified definition captured a significant proportion of biotechnology patents. It is extremely difficult to capture all biotechnology patents because the OECD definition of biotechnology is very broad. Generally speaking, the broader a technology field is, the more difficult it is to identify the corresponding classes, as they will be spread in different higher level categories, and possibly mixed with other technologies which are not of interest. For instance, although the list of classes for biotechnology in the table is concentrated in sections A, C and G of the IPC, certain patents (in the field) might be found in sections B, D and E but are mixed with other technology domains and cannot be separated (*e.g.* bioinformatics can be assigned to G06F but this class includes other computer-related technologies). Patent information based on the lists mentioned above therefore might not be complete, although the problem is not expected to be extensive.

As a method of validating the definition developed by selecting IPC codes, a further test was carried out using patent information from “biotechnology” firms. This process involved collecting individual patents owned by biotechnology firms and analysing the technology areas, as recorded by IPC codes. The names of the biotechnology firms were taken from the Finnish Bioindustries Website

[http://www.finbio.net/members/], which was then used to identify the patents owned by those firms. In the OECD database, 28 Finnish biotechnology firms were identified and statistical analysis on the primary and secondary IPC codes of the patents owned by those firms was conducted. The analysis of the sample of patents from Finnish "biotechnology firms" provided further evidence of the robustness of the definition. Unfortunately, it was not possible to carry out this validation test using biotechnology firms from other OECD countries due to lack of information.

It should be noted that, while the provisional definition appears to capture a significant proportion of biotechnology patents, it will also include some patents which are not for biotechnology techniques or products according to this Framework. These will mainly be found in classes under C07G (comprising compounds of unknown constitution: antibiotics, vitamins and hormones) and several classes under G01N (referring to a variety of techniques for investigating or analysing materials). However, the amount of error is likely to be small because the G01N class accounted for only 10.7% of all EPO 2000 applications from the IPC classes in Table 5 and there were no applications in the CO7G class.

Table 5. Provisional definition of biotechnology patents

IPC codes	Title
A01H 1/00	Processes for modifying genotypes
A01H 4/00	Plant reproduction by tissue culture techniques
A61K 38/00	Medicinal preparations containing peptides
A61K 39/00	Medicinal preparations containing antigens or antibodies
A61K 48/00	Medicinal preparations containing genetic material which is inserted into cells of the living body to treat genetic diseases; Gene therapy
C02F 3/34	Biological treatment of water, waste water, or sewage: characterised by the micro-organisms used
C07G 11/00	Compounds of unknown constitution: antibiotics
C07G 13/00	Compounds of unknown constitution: vitamins
C07G 15/00	Compounds of unknown constitution: hormones
C07K 4/00	Peptides having up to 20 amino acids in an undefined or only partially defined sequence; Derivatives thereof
C07K 14/00	Peptides having more than 20 amino acids; Gastrins; Somatostatins; Melanotropins; Derivatives thereof
C07K 16/00	Immunoglobulins, *e.g.* monoclonal or polyclonal antibodies
C07K 17/00	Carrier-bound or immobilised peptides; Preparation thereof
C07K 19/00	Hybrid peptides
C12M	Apparatus for enzymology or microbiology
C12N	Micro-organisms or enzymes; compositions thereof
C12P	Fermentation or enzyme-using processes to synthesise a desired chemical compound or composition or to separate optical isomers from a racemic mixture
C12Q	Measuring or testing processes involving enzymes or micro-organisms; compositions or test papers therefor; processes of preparing such compositions; condition-responsive control in microbiological or enzymological processes
C12S	Processes using enzymes or micro-organisms to liberate, separate or purify a pre-existing compound or composition processes using enzymes or micro-organisms to treat textiles or to clean solid surfaces of materials
G01N 27/327	Investigating or analysing materials by the use of electric, electro-chemical, or magnetic means: biochemical electrodes
G01N 33/53*	Investigating or analysing materials by specific methods not covered by the preceding groups: immunoassay; biospecific binding assay; materials therefore
G01N 33/54*	Investigating or analysing materials by specific methods not covered by the preceding groups: double or second antibody: with steric inhibition or signal modification: with an insoluble carrier for immobilising immunochemicals: the carrier being organic: synthetic resin: as water suspendable particles: with antigen or antibody attached to the carrier via a bridging agent: Carbohydrates: with antigen or antibody entrapped within the carrier
G01N 33/55*	Investigating or analysing materials by specific methods not covered by the preceding groups: the carrier being inorganic: Glass or silica: Metal or metal coated: the carrier being a biological cell or cell fragment: Red blood cell: Fixed or stabilised red blood cell: using kinetic measurement: using diffusion or migration of antigen or antibody: through a gel
G01N 33/57*	Investigating or analysing materials by specific methods not covered by the preceding groups: for venereal disease: for enzymes or isoenzymes: for cancer: for hepatitis: involving monoclonal antibodies: involving limulus lysate
G01N 33/68	Investigating or analysing materials by specific methods not covered by the preceding groups: involving proteins, peptides or amino acids
G01N 33/74	Investigating or analysing materials by specific methods not covered by the preceding groups: involving hormones
G01N 33/76	Investigating or analysing materials by specific methods not covered by the preceding groups: human chorionic gonadotropin
G01N 33/78	Investigating or analysing materials by specific methods not covered by the preceding groups: thyroid gland hormones
G01N 33/88	Investigating or analysing materials by specific methods not covered by the preceding groups: involving prostaglandins
G01N 33/92	Investigating or analysing materials by specific methods not covered by the preceding groups: involving lipids, *e.g.* cholesterol
* Those IPC codes also include subgroups up to one digit (0 or 1 digit). For example, in addition to the code G01N 33/53, the codes G01N 33/531, GO1N 33/532, etc. are included.	

Commodity classifications

Commodity classifications are needed for the derivation of statistical indicators about the production and trade in biotechnologically produced goods. As the classification needs to be applied to both domestic production and international trade statistics, the classifications need to be related to the standard international classifications used for those purposes *i.e.* the Central Product Classification (CPC) and the Harmonised System (HS). The CPC is, to an extent, based on the industrial classification, ISIC, and does not uniquely identify biotechnology produced goods. The HS classification is designed primarily for use by Customs officers who are responsible for implementing national customs systems. The classification does not differentiate between biotechnologically produced goods and those produced using other processes, as it is not possible for such officers to be able to differentiate between them. Thus neither of the classifications facilitate derivation of statistics about biotechnologically produced goods.

There are no available trade data that are precisely limited to well-defined biotechnology products. The best available data are from the US Census Bureau, which defines "biotechnology products" as a group that is almost entirely based on biologics. Biologics consists of therapeutic products derived directly from living organisms; these include vaccines, human blood and plasma, proteins and monoclonal antibodies. Major biotechnology drugs, such as humulin, interferon, epoetin, etc., fall under biologics. This definition both includes many products that are not part of advanced biotechnology and excludes other important biotechnologies. Nevertheless, this section follows the US Census practice in referring to "biotechnology" trade. Even though the US Census Bureau definition does not coincide with the definition adopted in this Framework, it is included here as an example of one use of international trade data to provide statistical indicators about biotechnology.

All of the biotechnology products on the Advanced Technology Products (ATP) list developed by the US Census Bureau[2] appear to belong to biologics. However not all biologics are derived from biotechnology, which means that the ATP commodity list – although at the ten-digit level – is still not adequately disaggregated to permit trade data on biotechnology alone. Furthermore, this definition of biotechnology as biologics excludes other products, such as scientific equipment used in biotechnology research, environmental biotechnology and agricultural biotechnology.

While the ATP list is an approximation only, it may be a useful starting point for discussing what constitutes an appropriate set of products. It is therefore presented with the understanding that further consideration and study is required. The list of commodities included in the biotechnology/biologics category is specified below.

The inclusions within the US Census Bureau definition of biotechnology imports are shown in Table 6 below.

2. www.census.gov/foreign-trade/www/sec2.html#ATP.

Table 6. Harmonized System categories of biotechnology commodities (ten-digit)

Harmonized Tariff Schedule commodity code

IMPORTS		
Year(s)	**Description**	
1994-2003	**2933294500**	**DRUGS (EXC AROM OR MOD AROM) CONT AN UNFUS IMI ETC.**
1994-2001	2937100000	PITUITARY (ANTERIOR) OR SIMILAR HORMONES
2002-2003	**2937110000**	**SOMATOTROPIN, ITS DERIVS & STRUCT ANALOGUES**
2002-2003	**2937190000**	**POLYPEPTIDE, PROTEIN & GLYCOPROTEIN HORMONES, NESOI**
2002-2003	**2937231010**	**ESTROGENS OF ANIMAL OR VEGETABLE ORIGIN**
2002-2003	**2937231050**	**PROGESTINS OF ANIMAL OR VEGETABLE ORIGIN, NESOI**
2002-2003	**2937235010**	**ESTROGENS NOT DERIV FROM ANIMAL OR VEGETABLE MATERIALS**
2002-2003	**2937235020**	**PROGESTERONE NOT DERIV FR ANIMAL OR VEGETABLE MATERIALS**
2002-2003	**2937235050**	**PROGESTINS NOT OF ANIMAL OR VEGETABLE ORIGIN, NESOI**
2002-2003	**2937399000**	**CATECHOLAMINE HORMONES, DERIVS & ANALOGUES NESOI**
2002-2003	**2937409000**	**HORMONE AMINO-ACID DERIVATIVES, NESOI**
2002-2003	**2937500000**	**PROSTAGLANDINS, THROMBOXANES & LEUKOTRIENES**
2002-2003	**2937900000**	**HORMONES, PROSTAGLANDINS, ETC. NESOI**
1994-2001	2937921010	ESTROGENS OF ANIMAL OR VEGETABLE ORIGIN
1994-2001	2937921050	OTHER PROGESTINS OF ANIMAL OR VEGETABLE ORIGIN
1994	2937924000	ETHYNODIOL DECANOATE; D-NORGRESTREL; AND DL-NORGRESTREL
1995-2001	2937925010	ESTROGENS NOT DERIV FROM ANIMAL OR VEGETABLE MATERIALS
1995-2001	2937925020	PROGESTERONE NOT DERIV FR ANIMAL OR VEGETABLE MATERIALS
1995-2001	2937925050	OTHER PROGESTINS NOT OF ANIMAL OR VEGETABLE ORIGIN
1994	2937928010	ESTROGENS NOT DERIVED FROM ANIMAL OR VEGETABLE MATERIALS
1994	2937928050	OTHER PROGESTINS NOT DERIVED FROM ANIMAL OR VEGETABLE MATERIALS
1994	2937994000	NANDROLONE DECANOATE; AND PIPECURONIUM BROMIDE
1994	2937998050	OTHER HORMONES AND THEIR DERIVATIVES, OTHER STEROIDS USED PRIMARILY AS HORMONES
1995-2001	2937999550	OTHER HORMONES AND DERIVATIVES, OTHER STEROIDS ETC.
1996-2003	**2940002000**	**D-ARABINOSE**
1996-2003	**2940006000**	**SUGARS, CHEM PURE (EXC SUCROSE, LACTOSE, ETC.) NESOI**
1995-2003	**3002200000**	**VACCINES FOR HUMAN MEDICINE**
1996-2003	**3002300000**	**VACCINES FOR VETERINARY MEDICINE**
1995	3002390000	OTHER VACCINES FOR VETERINARY MEDICINE
1995-2001	3002905050	OTH TOXINS MICROORGANISMS CULTURES & SIM PROD NESOI
2002-2003	**3002905150**	**BLOODS, VACCINES, TOXINS, ETC. NESOI**

The first column lists the year or range of years during which the HS code was valid. The second column lists the HS numbers, both current (in bold) and obsolete and a brief description. There were no HS changes in this area from 2003 to 2004**.** NESOI = not elsewhere specified or included.

The inclusions within the US Census Bureau definition of biotechnology exports are shown in Table 7 below.

Table 7. Harmonized System categories of biotechnology commodities (ten-digit)

Harmonized Tariff Schedule commodity code		
EXPORTS		
1995-2001	2937100000	PITUITARY (ANTERIOR) OR SIMILAR HORMONES
2002-2003	**2937110000**	**SOMATOTROPIN, ITS DERIVS & STRUCT ANALOGUES**
2002-2003	**2937190000**	**POLYPEPTIDE, PROTEIN & GLYCOPROTEIN HORMONES, NESOI**
2002-2003	**2937230000**	**ESTROGENS AND PROGESTINS**
1995-2001	2937920000	ESTROGENS AND PROGESTINS
1994-1995	2940000000	SUGARS, CHEM PURE (EXC SUCROSE, LACTOSE, ETC.)
1996-2003	**2940002000**	**D-ARABINOSE**
1996-2003	**2940006000**	**SUGARS, CHEM PURE (EXC SUCROSE, LACTOSE, ETC.) NESOI**
1995-2001	3002100040	FETAL BOVINE SERUM (FBS)
1994	3002100050	OTHER BLOOD FRACTIONS NOT ELSEWHERE SPEC
1995-2001	3002100060	OTHER BLOOD FRACTIONS NESOI
2002-2003	**3002100140**	**FETAL BOVINE SERUM (FBS)**
2002-2003	**3002100190**	**BLOOD FRACTIONS NESOI**
1994-2003	**3002200000**	**VACCINES FOR HUMAN MEDICINE**
1996-2003	**3002300000**	**VACCINES FOR VETERINARY MEDICINE**
1994-1995	3002310000	VACCINES AGAINST FOOT AND MOUTH DISEASE
1994-1995	3002390000	OTHER VACCINES FOR VETERINARY MEDICINE
1994-2001	3002905020	ANTIALLERGENIC PREPARATIONS
1994-2001	3002905050	TOXINS, CULTURES OF MICRO-ORGANISMS AND SIM PROD
2002-2003	**3002905120**	**ANTIALLERGENIC PREPARATIONS, NESOI**
2002-2003	**3002905150**	**HUMAN BLOOD; ANIMAL BLOOD PREPARED FOR THERAP, NESOI**

The first column lists the year or range of years during which the HS code was valid. The second column, lists the HS numbers, both current (in bold) and obsolete and a brief description. There were no HS changes in this area from 2003 to 2004. NESOI = not elsewhere specified or included.

Socio-economic objective classification

The *Frascati Manual*, which proposes standard practice for surveys on research and development, recognises the importance of a socio-economic objective (or purpose) classification. More work is needed to identify appropriate socio-economic categories for biotechnology. It is important to note that biotechnology is not necessarily an end objective in its own right – rather it is a way of achieving other objectives such as improved health, a cleaner environment or better food products.

Notwithstanding this point, it is important to recognise that funding for biotechnology research will often be dependent on the objective of that R&D. Hence, the existing *Frascati* classification should be applied to biotechnology R&D measured as part of a survey of biotechnology R&D performers. While the classification is used more frequently in R&D surveys of the government and higher education sectors, experience has shown that it can also be applied in the other sectors as well. Wherever possible the classification should be used in the publication of biotechnology R&D data.

Field of science and technology classification

The *Frascati Manual* for the measurement of R&D recognises the importance of a field of science and technology classification for the analysis of R&D data. However, biotechnology R&D covers many fields of research and so the classification cannot be used at a broad level to uniquely identify biotechnology. This finding was confirmed by recent Australian work (OECD, 2004*c*) showing that

biotechnology R&D can be reported within many fields of science and technology, at both the broad and detailed levels. The application of a more detailed classification may facilitate the derivation of biotechnology R&D but only if the classification is designed to identify biotechnology.

The current review of the *Frascati Manual* identified the need to examine the issue of the classification of fields of science and technology, leading to the formation of a separate task force to examine this issue. Future editions of the Biotechnology Framework will reflect the deliberations of that task force.

Application classification

Biotechnology techniques based on DNA/RNA or bioinformatics (see Box 1 of Chapter 2) have a wide range of applications. For example, DNA/RNA technologies have applications in human health, animal science, horticulture, environmental engineering, and production processes, while bioinformatics conceivably has many applications. As it is the application of biotechnology that determines its economic impacts, it is important for surveys of key biotechnology firms and research institutes (public and private) to obtain information on the area of application.

Table 8 provides a three-level classification system for applications. Although the broad level provides useful categories, the intermediate level will often provide a useful compromise that can increase the value of the results for policy applications. For example, using the two categories for agriculture provides essential information on the development and diffusion of GM versus non-GM applications of biotechnology to food and other cash crops. The detailed classification is likely to be most appropriate for surveys of specific sectors, such as the use of industrial bio-processes in manufacturing, where a breakdown at the two-digit sector level may be possible, or a survey of applications to human health.

Table 8. Proposed classification for biotechnology applications

Broad	Intermediate	Detailed
Human Health	Large molecule therapeutics and monoclonal antibodies (MABs) produced using rDNA technology	-
	Other therapeutics, artificial substrates, diagnostics and drug delivery technologies, etc.	Other therapeutics, drug delivery technologies, etc.
		Substrates (artificial bone, skin etc.)
		Diagnostics
Veterinary health	As above, for veterinary uses	As above
Agriculture	New varieties of genetically modified (GM) plants, animals, and micro-organisms for use in agriculture, aquaculture, and silviculture	GM plants, including fruit trees, flowers, horticultural crops, grains, etc.
		GM animals for agriculture
		GM fish
		GM tree varieties for forestry
		GM micro-organisms for agriculture (including bio pest control)
	New varieties of non-GM plants, animals, and micro-organisms for use in agriculture, aquaculture, silviculture, bio pest control and diagnostics **developed using biotechnology techniques** (DNA markers, tissue culture, etc.)	Non-GM plants, including fruit trees, flowers, horticultural crops, grains, etc.
		Non-GM animals for agriculture
		Non-GM fish
		Non-GM tree varieties for forestry
		Non-GM micro-organisms for agriculture (including bio pest control)
		Diagnostics
Natural resources	Applications for mining, petroleum/energy extraction, etc.	Mining: extraction using micro-organisms, etc.
		Petroleum/energy: extraction using micro-organisms
		Other resource applications
Environment	Diagnostics, soil bioremediation, treatment of water, air, and industrial effluents using micro-organisms, clean production processes	Diagnostics
		Soil bioremediation, including phytoremediation
		Effluent treatment
		Clean production processes
Industrial processing	Bioreactors to produce new products (chemicals, food, ethanol, plastics, etc.), biotechnologies to transform inputs (bioleaching, biopulping, etc.)	Detailed list of specific biotechnologies that are relevant to the firm's sector of activity[3]
Non-specific applications	Research tools, etc.	-
Other		-

3. The list of specific technologies would need to be regularly updated to reflect biotechnology applications in specific sectors.

CHAPTER 6: LINKS TO OTHER MANUALS

Further development of the Biotechnology Statistics Framework needs to be aligned closely with on-going work on the *Frascati Manual* (for R&D surveys). There is a lesser need to develop links with the *Oslo Manual* (for innovation surveys).

Alignment with the *Frascati Manual* is especially important in the following areas:

1. Specific indicators and definitions relating to R&D measures.

2. Specific elements of the institutional sector, industry and size and other classifications adopted in the *Frascati Manual*.

3. The socio-economic and field of science classifications – if possible at a detailed level to enable better identification of biotechnology R&D.

With respect to the *Oslo Manual*, it is recommended that future work consider development of suitable questions to enable the identification of biotechnology firms in future innovation surveys.

ANNEX 1: GLOSSARY OF TERMS USED IN THE LIST-BASED DEFINITION

The definitions below were provided by several delegates to the *Ad hoc* Biotechnology Statistics Group and were supplemented by the following sources:

- http://biotechterms.org/sourcebook/index.phtml
- http://www.bmbf.de/pub/systems_biology.pdf
- http://filebox.vt.edu/cals/cses/chagedor/glossary.html
- http://www.fao.org/biotech/index_glossary.asp?lang=en
- http://biotech.icmb.utexas.edu/search/dict-search.mhtml
- http://www.nbtc.cornell.edu/
- http://www.nanobioforum.org/files/article_BizInk_5Sep2003.pdf

DNA/RNA: Genomics, pharmacogenomics, gene probes, genetic engineering, DNA/RNA sequencing/synthesis/amplification, gene expression profiling, and use of antisense technology.

- Genomics/pharmacogenomics: The study of genes and their function. Advances in genomics due to the Human Genome Project and other genome research into plants, animals and micro-organisms are enhancing our understanding of the molecular mechanisms of genomes. Genomics stimulates the discovery of health care products by revealing thousands of new biological targets for the development of drugs and by identifying innovative ways to design new drugs, vaccines and DNA diagnostics. Genomic-based therapeutics includes both protein drugs and small molecule drugs. Genomics is also used in plant and animal breeding programmes.

- Gene probes/DNA markers: A section of DNA of known structure or function which is marked with a radioactive isotope, dye or enzyme so that it can be used to detect the presence of specific sequences of bases in another DNA or RNA molecule.

- Genetic engineering: Altering the genetic material of cells or organisms in order to make them capable of making new substances or performing new functions.

- DNA/RNA sequencing: Determination of the order of nucleotides (*i.e.* the base sequence) in a DNA or RNA molecule.

- DNA/RNA synthesis: The linking together of nucleotides to form DNA or RNA. In vivo, most synthesis involves DNA replication, but incorporation of precursors also occurs in repair. In the special case of retroviruses, an RNA template directs DNA synthesis.

- DNA/RNA amplification: The process of increasing the number of copies of a particular gene or gene-derived sequence.

- Other: There are several fields of research on RNA, including RNAi and siRNA, based on the use of recombinant technology to generate RNA sequences to inhibit gene function. Expression profiling analyses expressed genes using microarrays or gene chips.

Proteins and other molecules: Sequencing/synthesis/engineering of proteins and peptides (including large molecule hormones); improved delivery methods for large molecule drugs; proteomics, protein isolation and purification, signaling, identification of cell receptors.

- Peptide/Protein sequencing: Determination of the order of amino acids in a protein or peptide.

- Peptide synthesis: A procedure which links two or more amino acids in a linkage called a peptide bond.

- Protein engineering: The selective, deliberate (re)designing and synthesis of proteins. This is done in order to cause the resultant proteins to carry out desired (new) functions. Protein engineering is accomplished by changing or interchanging individual amino acids in a normal protein. This may be done via chemical synthesis or recombinant DNA technology (*i.e.* genetic engineering). "Protein engineers" (actually genetic engineers) use recombinant DNA technology to alter a particular nucleotide in the triplet codon of the DNA of a cell. In this way it is hoped that the resulting DNA codes for the different (new) amino acid in the desired location in the protein produced by that cell.

- Proteomics: Analysis of the expression, functions and interactions of all proteins of an organism.

- Signaling: Analysis of signaling molecules such as cytokines, chemokines, transcription factors, cell cycle proteins, and neurotransmitters.

- Cell receptors: Structures (typically proteins) found in the plasma membrane (surface) of cells that tightly bind specific molecules (organic molecules, proteins, viruses etc.). Some (relatively rare) receptors are located inside the cell (*e.g.* free-floating receptor for Retin-A). Both (membrane and internal) types of receptors are a functional part of information transmission (*i.e.* signalling) of the cell.

Cell and tissue culture and engineering: Cell/tissue culture, tissue engineering (including tissue scaffolds and biomedical engineering), cellular fusion, vaccine/immune stimulants, embryo manipulation.

- Cell/tissue/embryo culture and manipulation: Growth of cells, tissues or embryonic cells under laboratory conditions.

- Tissue engineering: Refers to the technologies used to induce:

  - (Injected) liver, cartilage, etc., cells to grow (within a recipient organism's body) and form replacement [integral] tissues.

  - (Extant) cells within the body encouraged to grow and form desired tissues, via precise injection of relevant compounds (*e.g.* certain growth factors, growth hormones, stem cells, etc.).

  - Laboratory grown tissue or organs to replace or support the function of defective or injured body parts (an example is skin tissue culture for grafts).

- Cell fusion: The combining of cell contents of two or more cells to become a single cell. Fertilisation is such a process.

- Vaccines/immune stimulants: A preparation containing an antigen consisting of whole disease-causing organisms (killed or weakened), or parts of such organism is used to confer immunity against the disease that the organisms cause. Vaccine preparations can be natural, synthetic or derived by recombinant DNA technology.

Process biotechnology techniques: Fermentation using bioreactors, bioprocessing, bioleaching, biopulping, biobleaching, biodesulphurisation, bioremediation, biofiltration and phytoremediation.

- Bioreactor: A vessel in which cells, cell extracts or enzymes carry out a biological reaction. Often refers to a fermentation vessel for cells or micro-organisms.

- Bioprocessing: A process in which living cells or components are used to produce a product, especially a biological product involving genetic engineering for commercial use.

- Bioleaching: The conversion of metals to a soluble form by live organisms such as bacteria or fungi.

- Biopulping: Use of micro-organisms to break down wood fibres for the purpose of producing pulp.

- Biobleaching: Use of micro-organisms to bleach pulp.

- Biodesulphurisation: Use of specific micro-organisms to transform hazardous sulphurs into less hazardous compounds.

- Bioremediation/biofiltration/phytoremediation: The process by which living organisms act to degrade hazardous organic contaminants or transform hazardous inorganic contaminants to environmentally safe levels in soils, subsurface materials, water, sludge, and residues.

  - Bioremediation: The use of micro-organisms to remedy environmental problems rendering hazardous wastes non-hazardous.

  - Biofiltration: The use of a support containing specific bacteria to capture by filtration hazardous substances from a gas stream.

  - Phytoremediation: Refers to the use of specific plants to remove contaminants or pollutants from either soils (*e.g.* polluted fields) or water resources (*e.g.* polluted lakes).

Gene and RNA vectors: Gene therapy, viral vectors.

- Gene therapy: Gene delivery, the insertion of genes (*e.g.* via retroviral vectors) into selected cells in the body in order to:

  - Cause those cells to produce specific therapeutic agents.

  - Cause those cells to become (more) susceptible to a conventional therapeutic agent that previously was ineffective against that particular condition/disease.

- Cause those cells to become less susceptible to a conventional therapeutic agent.
- Counter the effects of abnormal (damaged) tumour suppressor genes via insertion of normal tumour suppressor genes.
- Cause expression of ribozymes that cleave oncogenes (cancer-causing genes).
- Introduce other therapeutics into cells.

- Viral vectors: Certain (retro-) viruses that are used by genetic engineers to carry new genes into cells.

Bioinformatics: Construction of databases on genomes, protein sequences; modelling complex biological processes, including systems biology.

- The use of computers in solving information problems in the life sciences; mainly, it involves the creation of extensive electronic databases on genomes, protein sequences, etc. Secondarily, it involves techniques such as the three-dimensional modelling of **biomolecules.**

- The generation/creation, collection, storage (in databases), and efficient utilisation of data/information from genomics (functional genomics, structural genomics, etc.), combinatorial chemistry, high-throughput screening, proteomics, and DNA sequencing research efforts in order to accomplish a (research) objective (*e.g.* to discover a new pharmaceutical or a new herbicide, etc.). Examples of the data/information that is manipulated and stored include gene sequences, biological activity/function, pharmacological activity, biological structure, molecular structure, protein-protein interactions, and gene expression products/amounts/timing.

Nanobiotechnology: Applies the tools and processes of nano/microfabrication to build devices for studying biosystems and applications in drug delivery, diagnostics etc.

- Covers the interface between physics, biology, chemistry and the engineering sciences and which, among other things, aims to develop completely new measuring technologies for the biosciences.

Nanotechnology develops or makes materials that function on a very small scale, typically between 1 and 100 nanometers. Nanobiotechnology uses these particles and materials as tools to improve the performance and sensitivity of several life science technologies *e.g.* biosensing, medical devices and medical implants.

ANNEX 2: MODEL SURVEY OF BIOTECHNOLOGY USE AND DEVELOPMENT

This survey measures the activity of businesses involved in key biotechnology activities, that is, R&D into biotechnology techniques and the use of those techniques to develop products and processes or produce goods and services. It excludes end uses that involve widespread use of those products and processes in, for instance, manufacturing or simple environmental remediation using purchased biotechnology products. For more information on concepts, see Chapter 2.

The questions presented in the model questionnaire should be preceded by instructions per the normal practice of the surveying agency. Minimal definitions have been included in the questionnaire, and therefore it is recommended that agencies include definitions covering at least: general R&D concepts (per the *Frascati Manual*), employment and revenue. See Chapter 4 for more information on the survey.

Reference period/date: It is assumed that the reference period will be the most recent financial year and the reference date will be the last day of that year.

Location of firm activities: All questions in the model questionnaire are limited to the country where the firm is located. The restriction of firm activities to the country of location can either be made in general instructions to the respondent firm or in each question. It is possible to add questions that seek answers on activities outside the country of location, as long as these are clearly marked.

Scope of the model survey: The scope of the model survey is biotechnology active firms (enterprises), as defined in Chapter 2. These are firms involved in key biotechnology activities.

The target population of the model survey is discussed in Chapter 4. It is:

- Firms performing biotechnology R&D.
- Manufacturing firms currently undertaking key biotechnology activities.
- Firms classified to industries other than manufacturing or R&D if they are engaged in key biotechnology activities. These might include some firms classified to wholesaling.
- Those services firms that use process biotechnology techniques for the purpose of providing a service, for example, waste management or environmental remediation firms.

Firms to be excluded from the model survey are:

- Services firms that provide basic contract research services such as routine diagnostics and testing or firms providing consultancy and financial services to biotechnology firms.
- Biotechnology equipment and other goods suppliers that only distribute biotechnology products.
- End users of biotechnology products and processes, as described in Chapter 2.

There is no scope restriction based on employment size; see Chapter 5 of the Framework.

Important definitions: Members of the OECD's *Ad hoc* Biotechnology Statistics Group have developed a single and a list-based definition of biotechnology. It is important that both are used in surveys

by participating countries. The single definition is: *the application of science and technology to living organisms, as well as parts, products and models thereof, to alter living or non-living materials for the production of knowledge, goods and services.* The list-based definition can be found in Chapter 2 and is used in Question 1 of the model questionnaire (though note that the "other" category in the question is not part of the definition but is included following usual statistical practice). Definitions can be found in Chapter 2 for biotechnology R&D, products, processes, employment and revenue.

Methodology and data quality recommendations: See Chapter 4 of the Framework for a discussion of methodology (in particular, development of population lists) and data quality issues and recommendations.

Question 1 Biotechnology activities

Please indicate below your firm's activities in <period> for each of the listed biotechnology techniques.

If you are involved in activities not listed but which you regard as biotechnology, please provide details for them under "Other: please specify" at the end of the table.

Tick the applicable squares in each row.

Biotechnology	**During <period>, did your firm research or use this biotechnology?**	**During <period>, did your firm:**		
		do **research into** this biotechnology?	**use** this biotechnology for product or process development?	**use** this biotechnology in production (including for environmental purposes?)
DNA/RNA – genomics, pharmacogenomics, gene probes, genetic engineering, DNA/RNA sequencing/synthesis/amplification, gene expression profiling, and use of antisense technology.	☐ Yes → ☐ No	☐	☐	☐
Proteins and other molecules – sequencing/synthesis/engineering of proteins and peptides (including large molecule hormones); improved delivery methods for large molecule drugs; proteomics, protein isolation and purification, signaling, identification of cell receptors.	☐ Yes → ☐ No	☐	☐	☐
Cell and tissue culture and engineering – cell/tissue culture, tissue engineering (including tissue scaffolds and biomedical engineering), cellular fusion, vaccine/immune stimulants, embryo manipulation.	☐ Yes → ☐ No	☐	☐	☐
Process biotechnology techniques – fermentation using bioreactors, bioprocessing, bioleaching, biopulping, biobleaching, biodesulphurisation, bioremediation, biofiltration and phytoremediation.	☐ Yes → ☐ No	☐	☐	☐
Gene and RNA vectors - gene therapy, viral vectors.	☐ Yes → ☐ No	☐	☐	☐
Bioinformatics - construction of databases on genomes, protein sequences; modelling complex biological processes, including systems biology.	☐ Yes → ☐ No	☐	☐	☐
Nanobiotechnology – applies the tools and processes of nano/microfabrication to build devices for studying biosystems and applications in drug delivery, diagnostics etc.	☐ Yes → ☐ No	☐	☐	☐
Other: please specify _____ _____	☐ Yes → ☐ No	☐	☐	☐
Other: please specify _____ _____	☐ Yes → ☐ No	☐	☐	☐

Did you answer Yes in column 2 to any of the listed techniques?

☐ → No please go to Question 8 ☐ → Yes please go to Question 2

Question 2 Biotechnology products and strategy

A **biotechnology product** can be a good or service. Its development required the use of one or more of the biotechnologies listed in Question 1. A **biotechnology process** is defined as a production or other process using one or more biotechnology techniques or products.

	Yes	No
At <date> did your firm have biotechnology products on the market?	☐	☐
Is your firm currently developing products that require the use of biotechnology?	☐	☐
Is your firm currently developing processes that require the use of biotechnology?	☐	☐
Do you consider that biotechnology is central to your firm's activities or strategies?	☐	☐

Question 3 Employment

General instructions

Report the number of persons employed in your firm at the last pay period of <reference period>.

In a, b and c below, count persons who are employed on a part-time basis as one person.

In d below, full-time equivalents (FTE) are defined as: One FTE may be thought of as one person-year. Thus, a person who normally spends 30% of his/her time on R&D and the rest on other activities (such as teaching, university administration and student counselling) should be considered as 0.3 FTE. Similarly, if a full-time R&D worker is employed at an R&D unit for only six months, this results in an FTE of 0.5. Since the normal working day (period) may differ from sector to sector and even from institution to institution, it is not meaningful to express FTE in person-hours (Frascati Manual).

Include*: working proprietors and partners.*

Exclude*: consultants and contractors who are not employees and unpaid staff e.g. student volunteers.*

Persons employed on biotechnology activities (as reported in Question 1)

Include*: researchers, managers, production workers, and support staff who are directly involved in biotechnology activities.*

Exclude*: indirect support staff such as central personnel and central IT staff.*

a How many persons, in total, worked for your firm at <date>?	☐
b Number of biotechnology employees: persons who worked for your firm at <date> who spent some or all of their time on biotechnology activities during <period>	☐
c Of your biotechnology employees (reported in b above), how many were:	
Primarily engaged in biotechnology R&D……	☐
Primarily engaged in other biotechnology activities (*e.g.* production)……..	☐
d Please estimate the total full-time equivalents (FTEs) in biotechnology	
Total FTE spent in the performance of biotechnology R&D……..	☐
Total FTE spent in the performance of other biotechnology activities (*e.g.* production)	☐

Question 4 Status of biotechnology applications

Please indicate the status of your firm's biotechnology activities for each of the following applications (as at <reference date>).

Tick the applicable squares in each row.

Biotechnology application	R&D	Pre-clinical trials/ confined field trials	Regulatory phase/ unconfined release assessment	Approved/ marketed /in production	Not relevant
Human health – large molecule therapeutics and monoclonal antibodies produced using rDNA technology	☐	☐	☐	☐	☐
Human health – other therapeutics, artificial substrates, diagnostics and drug delivery technology etc.	☐	☐	☐	☐	☐
Veterinary health – all health applications for animals	☐	☐	☐	☐	☐
GM agricultural biotechnology – new varieties of genetically modified (GM) plants, animals and micro-organisms for use in agriculture, aquaculture, and silviculture	☐	☐	☐	☐	☐
Non-GM agricultural biotechnology – New varieties of non-GM plants, animals, and micro-organisms for use in agriculture, aquaculture, silviculture, bio pest control and diagnostics **developed using biotechnology techniques** (DNA markers, tissue culture, etc.)	☐	☐	☐	☐	☐
Natural resource extraction – applications for mining, petroleum/energy extraction, etc.	☐	☐	☐	☐	☐
Environment – diagnostics, soil bioremediation, treatment of water, air, and industrial effluents using micro-organisms, clean production processes	☐	☐	☐	☐	☐
Industrial processing – bioreactors to produce new products (chemicals, food, ethanol, plastics, etc.), biotechnologies to transform inputs (bioleaching, biopulping, etc.)	☐	☐	☐	☐	☐
Non-specific applications – research tools etc.	☐	☐	☐	☐	☐
Other: please specify _____ _____	☐	☐	☐	☐	☐
Other: please specify _____ _____	☐	☐	☐	☐	☐

Question 5 Financial characteristics	
Total value of firm sales/revenues from all sources for <period>……..	☐
Value of revenues from biotechnology activities for <period>……..	☐
Total R&D spending within the firm for <period>……..	☐
Total R&D spending on biotechnology activities within the firm for <period>……..	☐
How much venture capital did your firm raise for biotechnology activities in <period>?.....	☐

Question 6 Intellectual property protection	
In respect of the biotechnology activities reported in Question 1, how many <u>technically unique*</u> biotechnology patents does your firm have?	
already granted/approved as at <date>?	☐
applied for during <reference period>?	☐
* Do not double count patents in more than one jurisdiction for the same invention.	

Question 7 Barriers to biotechnology R&D and commercialisation

Which of the following factors were <u>significant</u> barriers to your firm's biotechnology R&D activities or your ability to commercialise biotechnology products?

Tick the applicable squares in each column.

	R&D	**Commercialisation**
Access to capital	☐	☐
Access to technology/information	☐	☐
Access to skilled human resources	☐	☐
Access to international markets	☐	☐
Lack of distribution and marketing channels	☐	☐
Public perception/acceptance	☐	☐
Regulatory requirements	☐	☐
Patent rights held by others/high licensing costs	☐	☐

Question 8 Comments
Thank you for completing this questionnaire. Please provide comments below on any of the information you have supplied or if you have any suggestions for improvement. …… …… …… …… …… ……

Supplementary information about the questionnaire

Question 1 has three purposes as follows:

- It assesses the biotechnology techniques being used/researched by performing firms and the type of activity (R&D or production).

- It identifies firms undertaking key biotechnology activities according to this Framework and screens other firms from the remainder of the questionnaire.

- It establishes the meaning of the term "biotechnology activities" as used in the remainder of the questionnaire.

Question 2 provides information on whether or not the responding firm has a biotechnology product on the market and whether biotechnology is strategically important for it.

Question 3 asks about biotechnology employment, split into persons employed on biotechnology R&D and other biotechnology activities. It also asks about FTE for the same split. Note the departure from the *Frascati Manual* in specifically excluding all consultants and contractors who are not employees of the firm (*Frascati* recommends the inclusion of on-site consultants and contractors in the measure of human resources).

An alternative for Question 2 includes adding a gender split against some or all of the employment questions.

Question 4 collects information about the areas of application of the firm's biotechnology activities and the status of those activities.

Question 5 collects financial information attributable to biotechnology activities. Note that only intramural R&D expenditure is asked for in this question, although some countries might like to also collect extramural expenditure.

Question 6 seeks the number of biotechnology patents that the firm has already had granted/approved and the number applied for during the reference period. The question should be related back to the list-based definition, rather than the classification presented in Chapter 5. Note that this question does not cover all of the patents which a firm might have (as it excludes those which have been pending since before the start of the reference period).

Question 7 collects information on barriers to biotechnology R&D and commercialisation.

There are a number of questions on highly relevant topics that participating countries have used but which have not been included in this version of the model questionnaire. See Chapter 4 (the section on *Statistics of interest which are not covered by this Framework*) for a brief discussion of such questions.

BIBLIOGRAPHY

ARUNDEL, A. (2002), "Agro-biotechnology, innovation and employment", *Science and Public Policy* 29:297-306.

ARUNDEL, A. (2003), "Biotechnology Indicators and Public Policy", *STI Working Papers 2003/5*, OECD, Paris, June.

ARUNDEL, A and B. Mintzes (2004), "The benefits of biopharmaceuticals", *Innogen Working Paper No. 14*, Edinburgh University, August.

DEVLIN, A (2003), "An Overview of Biotechnology Statistics in Selected Countries", *STI Working Papers 2003/13*, OECD, Paris, October.

OECD (1997), *Oslo Manual: Proposed Guidelines for Collecting and Interpreting Technological Innovation Data*, OECD, Paris.

OECD (2002a), *Frascati Manual, Proposed Standard Practice for Surveys on Research and Experimental Development*, OECD, Paris.

OECD (2002b), "A Statistical Framework for Biotechnology Statistics", internal working document, OECD, Paris.

OECD (2002c), "Biotechnology Use and Development: Towards A Model Survey", internal working document, Paris.

OECD (2003), *Quality Framework and Guidelines for OECD Statistical Activities*, http://www.oecd.org/dataoecd/26/42/21688835.pdf, OECD, Paris.

OECD (2004*a*), "Quality Indicators", internal working document, OECD, Paris.

OECD (2004*b*), "Construction of Biotechnology Survey Population Lists: Methods and Challenges", internal working document, OECD, Paris.

OECD (2004*c*), "Australia: Identification of Fields of Science Relevant to Biotechnology R&D", internal working document, OECD, Paris.

ROSE, A. (2000), "A Challenge for Measuring Biotechnology Activities", published in *The Economics and Social Dynamics of Biotechnology*, 2000.

SENKER, J. and P. VAN ZWANENBERG (2000), "European Biotechnology Innovation System: EC Policy Overview", DG Research, European Commission, TSER Contract No. SOEI-CT98-1117, September.

STATISTICS NEW ZEALAND (2003), *Research and Development in New Zealand 2002*, September.

TRAORE, Namatié (2004), "Biotechnology Use and Development Survey: Methodology, Issues and Responses, *Statistics Canada Working Paper 88F0006XIE — No. 006*, February.

VAN BEUZEKOM, Brigitte (2000), "Biotechnology Statistics in OECD Member Countries: An Inventory", *STI Working Paper 2000/6*, OECD, Paris, September.

VAN BEUZEKOM, Brigitte (2001), "Biotechnology Statistics in OECD Member Countries: Compendium of Existing National Statistics", *STI Working Paper 2001/6*, OECD, Paris, September.

VAN BEUZEKOM, Brigitte (2004), "Biotechnology Statistics in OECD Member Countries: An Inventory", *STI Working Paper 2004/8*, OECD, Paris, September.