



## EUROPE

CHILDREN AND FAMILIES  
EDUCATION AND THE ARTS  
ENERGY AND ENVIRONMENT  
HEALTH AND HEALTH CARE  
INFRASTRUCTURE AND  
TRANSPORTATION  
INTERNATIONAL AFFAIRS  
LAW AND BUSINESS  
NATIONAL SECURITY  
POPULATION AND AGING  
PUBLIC SAFETY  
SCIENCE AND TECHNOLOGY  
TERRORISM AND  
HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from [www.rand.org](http://www.rand.org) as a public service of the RAND Corporation.

Skip all front matter: [Jump to Page 1](#) ▼

### Support RAND

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

### For More Information

Visit RAND at [www.rand.org](http://www.rand.org)

Explore [RAND Europe](#)

View [document details](#)

### Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND Web site is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

# TECHNICAL REPORT

---

## Data strategies for policymaking

Identifying international  
good practice

Final report

Philipp-Bastian Brutscher, Jan Tiessen,  
Ala'a Shehabi, Daniel Schweppenstedde,  
Claire Celia, Christian van Stolk, Tom Ling

Prepared for the European Commission Directorate General for  
Health and Consumers (DG SANCO)

The research described in this report was prepared for the European Commission. The opinions expressed in this study are those of the authors and do not necessarily reflect the views of the European Commission.

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND**® is a registered trademark.

© Copyright 2010 European Commission

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the European Commission.

Published 2010 by the RAND Corporation  
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138  
1200 South Hayes Street, Arlington, VA 22202-5050  
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665  
Westbrook Centre, Milton Road, Cambridge CB4 1YG, United Kingdom  
RAND URL: <http://www.rand.org>  
RAND Europe URL: <http://www.rand.org/randeurope>  
To order RAND documents or to obtain additional information, contact  
Distribution Services: Telephone: (310) 451-7002;  
Fax: (310) 451-6915; Email: [order@rand.org](mailto:order@rand.org)

# Executive Summary

---

## Background and rationale

In recent years the European Commission has gradually moved towards a more evidence-based policymaking approach. The underlying assumption is that policy decisions based on systematic evidence produce better outcomes.

With the move towards evidence-based policymaking, the need for high-quality data – which can include statistical and qualitative data, but also existing research – has increased. The data needs for the increasingly formalised and institutionalised impact-assessment process are an example of how data needs of modern policymaking have increased.

DG SANCO has already done much to respond to this increased need – including developing its Consumer Markets Scorecard, its health economics database and its health statistics database. Despite these efforts, DG SANCO feels it needs to explore the possibility of having a more coordinated data strategy in order to be able to put the best evidence from research at the heart of policy development, implementation and enforcement.

## Developing a conceptual framework for data strategies

Drawing on the literature and 13 case studies of data strategies in different organisations dealing with data (for an overview of the organisations, see Table 0.1 below), we developed a simple framework that allows DG SANCO to think about its own data strategy in a systematic way. We discussed the corresponding key elements and drew out possible lessons for DG SANCO.

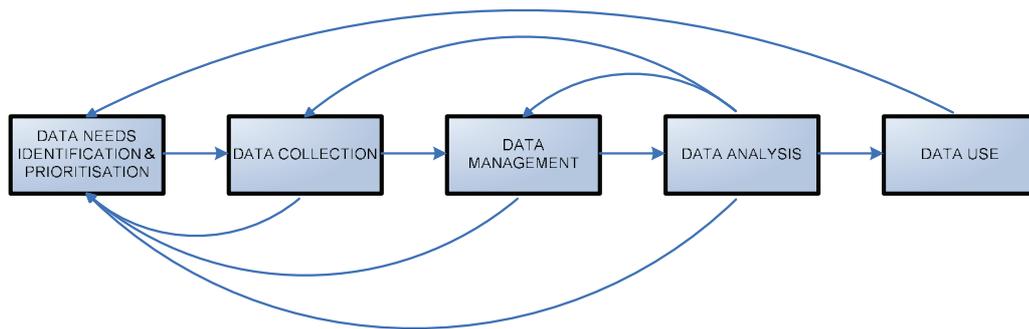
**Table 0.1: Overview of case studies**

	<b>Comparator</b>	<b>Country</b>	<b>Policy area</b>
1	Department for Environment, Food and Rural Affairs (DEFRA)	UK	Animal health
2	DG Enterprise	EU	General
3	DG Environment	EU	General
4	DG Trade	EU	General
5	Direction générale de la concurrence, de la consommation et de la répression des fraudes (DGCCRF)	France	Consumer
6	European Health Observatory (EHO)	International organisation	Public health
7	Food Standards Agency (FSA)	UK	Consumer
8	Instituto Nacional del Consumo (INC)	Spain	Consumer
9	Konsumentverket, Swedish Agency for Consumer Affairs (SCA)	Sweden	Consumer
10	Organisation for Economic Cooperation and Development (OECD)	International organisation	General
11	Public Health Agency Canada (PHAC)	Canada	Public health
12	Regulatory Reform Group (joint organisation of the ministries of Finance and Economic Affairs) (RRG)	Netherlands	General
13	World Health Organisation (WHO)	International organisation	Public health

The framework we developed represents an idealised, sequential process of developing a data strategy ‘from scratch’. It comprises five key elements (for an illustration see Figure 0.1: below):

1. identification and prioritisation of data needs
2. data collection
3. data management
4. data analysis and communication
5. data use.

The simple idea underlying the data-strategy framework is that the identification and prioritisation of data needs feeds into and determines the collection of data; which in turn feeds into data management (including data storage, verification and compilation); which again feeds into data analysis; which finally informs policymakers and so feeds into use of the data. At the same time we do expect there to be feedback mechanisms that do not follow the sequential model, but link different elements directly.



**Figure 0.1: Data-strategy framework**

## Lessons for DG SANCO

Using the data-strategy framework developed in the initial stage of the research, the key lessons learned and the recommendations for DG SANCO may be summarised as follows.

### Identifying and prioritising data needs

There are several ways of identifying data needs. We distinguished between decentralised and centralised approaches to data identification and discussed the advantages and disadvantages of a systematic approach to identifying data needs, as opposed to an ad hoc approach.

When it comes to the prioritisation of data needs, we found that organisations differ in terms of the degree of formality with which this is done. Some organisations do not prioritise their identified data needs at all; others do so in a very informal way; still others use a formal process. Formal approaches, in turn, may be categorised into top-down and bottom-up approaches.

#### *Lessons for DG SANCO*

The main lessons for DG SANCO regarding the identification of data needs that follow from our discussion include the following:

1. Data needs should – where possible – be identified in a systematic way, using appropriate methods to overcome the cognitive limitations that can often act as a barrier to the identification of data needs.
2. To the extent that subject-related and methodological expertise sits with individual policy units, a decentralised system of identifying data needs is generally preferable to a centralised one.
3. A decentralised system of identifying data needs can be complemented with a top-down way of prioritising data needs, which ensures high-level buy-in to empirical work. At the same time, it allows the collection of the data needed to answer the most pressing policy questions.

### Collecting data

When it comes to collecting data, we distinguished between organisations that collect data internally, externally and by means of a downstream agency; and organisations that collaborate with other organisations in data collection.

We discussed some of the pros and cons (in terms of costs and timeliness) of organising data collection in one of these ways. With regard to the question of what data to collect, or where to collect them from, we distinguished between primary data collection, a hybrid type of data collection (which relies on Member States and/or agencies to collect data) and secondary data collection.

We found that the hybrid form of data collection is the most popular of these – in particular if, as is often the case, primary data collection is not available. Regardless of which approach is chosen, basically all organisations stress the importance of having quality controls in place to organise the data collection. The main approaches used include peer review, pilot studies and transparency with regard to the collection approach, coverage, response rates and other essential characteristics of the collection process.

#### *Lessons for DG SANCO*

The lessons for DG SANCO regarding the collection of data that follow from our discussion include the following:

4. The organisation of data collection should take costs into account and rely primarily on collaborative collection efforts. Only where this is not possible on external (internal) data collection efforts if the same type of data is collected with low (high) frequency.
5. Similarly, it is advisable to rely on secondary data collection when possible and to collect primary data only where secondary data are not available/suitable. At the same time, the case study on the UK Department for Environment, Food, and Rural Affairs (DEFRA) illustrates that generally fewer relationships / secondary data sources are preferable to many.
6. Probably the most promising approach to assure the quality of data to be collected is to use a peer-review process at the planning and possibly the pilot stage of the data collection. There is good reason to try to involve people responsible for the data analysis in this process.

#### **Data management**

In the context of data management, we reviewed the advantages and disadvantages of a centralised/decentralised data-storage system and compared how data verification is handled across the organisations we reviewed.

In addition, we discussed the role of data compilation in the context of the work of different organisations and outlined the main strategies of dealing with identifiable data: not to collect identifiable data in the first place, or, if identifiable data is collected, to make sure that the data are anonymised before being disseminated and that the requestor of the data is legally bound to use the data sensibly.

#### *Lessons for DG SANCO*

From our discussion, again several interesting points for DG SANCO emerge. These include the following:

7. DG SANCO could use three main criteria to evaluate whether a centralised or decentralised data-storage system is more appropriate: i) Is there an overlap in the data needs of different policy units; ii) Are data needs rather limited in scope? iii) How timely do the data have to be?

8. Owing to the danger of encountering severe bias if data verification is designed in the wrong way, DG SANCO could benefit from a formal system for this, which could be done either in-house or by out-sourcing.
9. To the extent that DG SANCO intends to use its data for research purposes, there is good reason to take data compilation seriously. WHO, PHAC and OECD offer good examples of how data compilation can be organised (see Box 8).
10. Given the high risk associated with dealing with identifiable data, DG SANCO may want to adopt a policy of collecting data at the lowest level of identification possible, distributing identifiable data only in exceptional cases.

### Data analysis

For the analysis stage of data strategies, we reviewed and compared the different organisational approaches to data analysis in terms of who carries out the analysis, how they do so, and how the analysis is quality assured and communicated.

The main distinctions across the organisations in analysing data were the size and capacity of the organisation to carry out data analysis, the cost of having in-house analytical research competencies versus outsourcing, and benefits that are accrued to the organisation if analysis is done in-house (eg the accumulation of in-house expertise and streamlining the data strategy from beginning to end).

#### *Lessons for DG SANCO*

For the data-analysis stage of a data strategy, DG SANCO could take heed of the following points:

11. DG SANCO could evaluate its internal capability to carry out data analysis, identify where there are gaps in its knowledge, compare the costs of expanding its analytical capabilities relative to the cost of outsourcing / using external contractors, and measure the value added and the benefits of having in-house expertise in data analysis.
12. The merits of a centralised or decentralised approach to data analysis need to be assessed against DG SANCO's existing organisational structure. Good practice examples for either way of organising data analysis are EHO and DG Enterprise.
13. Good practice in quality assurance for data analysis involves transparency and communication of methodologies (principles, assumptions and methods) and systematic reviews of methodologies. OECD and PHAC provide standards that could be adopted by DG SANCO.

### Data use

In terms of data use, as the final element of our data-strategy framework, we found that there are four main users of data and analysis: the general public, government and public bodies/agencies, internal clients and academics.

These organisations use data for four main purposes: to inform policy; to monitor, review and evaluate; for academic use; and for internal use.

*Lessons for DG SANCO*

14. DG SANCO may want to consider data use and communication as part of an integrated data strategy and;
15. Identify the main users of its data-analysis products to ensure the wide and effective use and uptake of DG SANCO work
16. Understand how and for what purpose the end users use the analysis – eg by using SANCO internal workshops
17. Develop mechanisms to reach and involve users – tested examples include Scenario Logic and Scenario Gaming workshops.

**Revising the linear model**

Whilst in the early chapters we implicitly assume a linear model of data strategies – moving from identifying and prioritising data needs to collecting data, managing data, analysing data and using it – in the final chapter of this report we explicitly discuss some of the nonlinearities involved in data strategies.

More specifically, we argue that all (later) stages of a data strategy have bearing on the identification and prioritisation of data needs; that data management and analysis influence data collection; that data analysis influences data management; and that data use has some bearing on the analysis stage of data.

*Lessons for DG SANCO*

Several lessons for DG SANCO follow from our discussion in this chapter:

18. DG SANCO may want to structure its identification and prioritisation of data needs into three parts: i) identification of data needs; ii) feasibility; and iii) prioritisation. Policy analysts would be involved in part i), people responsible for collection and management in part ii) and senior management in part iii).
19. A formal system of workshops that allow people from the different stages to interact with one another – in particular those from data collection, management and analysis and data analysis and management – could be beneficial to DG SANCO.