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.
Evaluation of DG SANCO data management practices

Final report

Jan Tiessen, Claire Celia, Tom Ling, Helen Ridsdale, Maiwênn Bareaud, Christian van Stolk

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.
Executive summary

This study assesses DG SANCO’s current data management practice
RAND Europe was asked by DG SANCO to assess the current data management practice within DG SANCO. Data management in this definition comprises the use of data in policy making in its wider meaning. This study maps current practice, identifies perceived strength and weaknesses and develops recommendations to address them. The context of this work is the desire of DG SANCO to develop a comprehensive data strategy to meet the increasing needs of evidence-based policy making. Such a strategy would seek to establish good practice in how data are identified, collected, stored, analysed, used and communicated.

The findings of this report were based mostly on interviews with DG SANCO staff
The analysis contained in this report relies to a large extent on key informant interviews and thus resembles a perception audit of what a subset of DG SANCO staff see as current practice, its strengths and its weaknesses. In addition internal documentation obtained from DG SANCO was reviewed, and the research team analysed a sample of documents produced during the different stages of the policy process. This report focuses mostly on areas of improvement, though it was clear from the interview findings that current practice in DG SANCO had many strengths. Therefore, DG SANCO starts from a position of strength in data management. In other areas more substantial changes are required, in the opinion of staff.

DG SANCO’s current data practice is characterised by diversity between directorates and policy stages
DG SANCO staff use a relatively wide definition of data, including not only quantitative statistical data, but also more qualitative sources of data. The data needs between directorates differ substantially in terms of type and content of data; in general there appears to be little overlap in data needs between directorates, often even between units. Both (perceived) data needs and actual use of data differ substantially between policy stages. The impact assessment stage is currently the most data-heavy stage of the policy process, using the widest range of data types and sources; while enforcement and evaluation are the stages with the most comprehensive use of statistical information, often collected by DG SANCO or submitted by Member States.

DG SANCO produces some of its own data, but in the early stages of the policy process it mostly relies on the use of external data sources such as Eurostat, and the use of external contractors is widespread. For enforcement purposes, however, DG SANCO mostly relies
on self-generated data obtained through alert systems (e.g. RAPEX) or reporting requirements for Member States.

DG SANCO has access to a substantial amount of data that it requests from Member States, but these most of these data are deposited in a large number of different databases or even not stored in an electronic format. Overall, data storage is decentralised and fragmented. Data sharing and collaboration are taking place between colleagues at DG SANCO informally but there is a lack of formal knowledge about what staff are doing across units and directorates.

Data analysis tends to be descriptive and qualitative rather than statistical in nature, assessing context and problems rather than evaluating and projecting. There is a lack of quantitative data analysis and in-house specialist skills (i.e. those of economists and statisticians) to deal with data outside the specialist units, which is partially compensated for by the widespread use of contractors to analyse data.

DG SANCO staff currently conduct, for the most part, pragmatic assessments of the quality of data they use. The two main means by which staff assess data quality are (1) checking that the data source is reliable and (2) assessing the comparability of the data. There are no formal guidelines on how to assess the quality of data.

Perceived strengths and weaknesses vary by directorate and function

The perception of current strengths and weaknesses in using data for policy making differs between directorates and experience in data use, with staff from information units and Directorates D, E and F the most positive about current strengths and weaknesses. This points to pockets of good practice that should be extended to improve the overall performance of DG SANCO. Key issues mentioned both as strengths and weaknesses include the following:

- In general, collaboration and knowledge sharing between individuals was perceived as a strength and as working very well. There were, however, concerns that this collaboration was personalised at an individual level, which could be a barrier for new staff.

- Specialist skills to collect, analyse and use data are perceived to be in short supply at DG SANCO in general, but some units in Directorates D, E and F and the information units (B1, C2) have substantial skills for and experience in using quantitative data.

- While staff working on enforcement and in the information units were generally satisfied with the availability of data and consider it a strength, other respondents highlighted problems in access to data and knowledge about potential data sources.

- Nevertheless, respondents also felt that some of the data available to DG SANCO are not sufficiently used because staff are unaware of the existence of those data, or because quality and comparability make them difficult to use.

- Staff felt that there is the importance of data for DG SANCO is increasing, but the use of data is still very much ad hoc and data do not yet play a sufficient role in policy development;
The report includes strategic high-level and specific recommendations

High-level recommendations focus mostly on the shift in organisation and culture which will be required for the development of a comprehensive data strategy. Therefore, they are likely to represent changes that it may be difficult to implement and that require a long time horizon to embed. They focus on the following.

1. Clarify the purpose of a data strategy
   Prior to developing a high-level data strategy, DG SANCO should define and clarify the purpose of a data strategy and define what it wants data to be used for. Currently a wide range of sometimes competing purposes is mentioned by DG SANCO staff, ranging from monitoring enforcement action and implementation to communication to the general public. If DG SANCO wants to cover a wider range of purposes, it should make sure that the data strategy acknowledges the different uses for which data are intended.

2. Think ‘information’ and ‘knowledge’ rather than data
   In developing a comprehensive data strategy, DG SANCO should think about the knowledge and information it requires for policy making. Data will only become useful when structured, interpreted and understood in context. Thus DG SANCO should ensure that the capacity and skills to make use of data are available.

3. Prioritise
   Given the limited resources and skills available, it will be essential to prioritise the collection of specific data. This prioritisation should be based on the key questions that DG SANCO needs to answer – that is, the information and knowledge DG SANCO wants to obtain. For example, one approach that can be used is logic modelling, establishing a logic and theory of change.

4. Create a decentralised, but coordinated, organisational structure system
   Given the differences between directorates, both in term of subject area and type of activity, DG SANCO should consider a decentralised, but coordinated, organisational basis for its data strategy. This should be supplemented by shared resources for the whole organisation, and regular communication between these units should take place, including updates on their recent activities. Information units available should be used to a greater extent and could form the core of directorate-wide centres of expertise and support for policy units.

5. Aim for cultural change
   Any organisational and procedural change suggested here should be embedded in a process of cultural change. DG SANCO staff should feel that using data in their day-to-day work is relevant, valued and recognised within the wider organisation. Ultimately, DG SANCO staff will need to be aware that using data makes a substantial difference in the policy process by improving decisions taken.

Specific recommendations
Specific recommendations are closely linked to the weaknesses identified by staff, and this relationship is shown in Table 0.1 below.
Table 0.1.1: Linking recommendations to the issues identified

<table>
<thead>
<tr>
<th>Weakness mentioned</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of specialist skills in-house</td>
<td>Produce a map of data management specific skills of DG SANCO staff</td>
</tr>
<tr>
<td></td>
<td>Provide basic training on data use and management for DG SANCO staff</td>
</tr>
<tr>
<td>Some of the data within DG SANCO are either not used at all or not well used</td>
<td>Establish a central survey unit</td>
</tr>
<tr>
<td>Harmonisation issues with data obtained from Member States</td>
<td>Conduct a data inventory</td>
</tr>
<tr>
<td>Current data management practices are reactive and crisis driven</td>
<td>Create a central knowledge base or information hub on available data sources, both internal and external</td>
</tr>
<tr>
<td>Access to data is sometimes an issue, as well as how data are used</td>
<td>Develop methods to define data needs systematically</td>
</tr>
<tr>
<td>Policy is not sufficiently data led</td>
<td>Develop basic guidance on the use of data in day-to-day policy work</td>
</tr>
<tr>
<td></td>
<td>Develop guidance on how to assess and ensure data quality</td>
</tr>
<tr>
<td>Timeliness is an issue in data management</td>
<td>Plan and discuss data needs as early as possible in the policy making process.</td>
</tr>
<tr>
<td>Problems with collaboration in data collection and analysis</td>
<td>Make data use more visible.</td>
</tr>
<tr>
<td>No unified approach to data across DG SANCO</td>
<td>Be prepared for ad-hoc data needs.</td>
</tr>
<tr>
<td></td>
<td>Become more concise in contracting out research</td>
</tr>
<tr>
<td></td>
<td>Strengthen external collaboration on data collection</td>
</tr>
<tr>
<td></td>
<td>Create decentralised support structures</td>
</tr>
</tbody>
</table>

This report provides additional tools to implement a data strategy

The appendices to this report contain additional tools for DG SANCO to develop and implement a comprehensive data strategy, including a reflection on how to develop the high-level macroindicators that DG SANCO should use, some practical tips for staff, and indicators to help understand the effectiveness of a data management strategy. For the high-level indicators, the study proposes using a ‘theory of change’ approach, whereby DG SANCO examines its logics of intervention or how it produces impacts. The indicators are then developed on the basis of how DG SANCO achieves these impacts. In terms of practical tips, the report identified a number of key diagnostic questions for use by DG SANCO staff at each stage of the data management framework. These were identified on the basis of the interviews and offer a checklist for staff to help them improve data management. Finally, the impact of a data strategy has to be evaluated once changes have been made. This requires DG SANCO to establish a baseline and then look *ex-post* at how data management has become embedded and what its impact has been.