

Creating new statistics on a fuzzy concept

Camilla Østerballe Nielsen, Statistics Denmark, cnl@dst.dk

Lisbeth Harbo, Statistics Denmark, lhb@dst.dk

Abstract.

This paper presents aspects of the development of a new register on relatives by Statistics Denmark. Funded by TrygFonden, the wish for such a register on relatives originates in a current Danish debate on the potential caregiving resources – and the experienced strain – of relatives to people in need of care. The information in the register is based on Danish administrative data and is primarily to be used for research purposes, but forms also the basis for statistics on relatives in Statistics Denmark's online StatBank.

The funding organisation stipulated that the development of the statistics should be conducted in cooperation with the civil society within the social and health sectors who concern themselves with the topic of relatives and the role of caregiving. This is very much in line with the approach outlined in the GSBPM phase 1 where initial engagement of stakeholders in the development of new statistics leads to more relevant outcomes.

Engaging external stakeholders in the development process requires balancing the stakeholders' high expectations and wishes for the final outcome with the restrictions due to data availability and the framework for statistical production offered by Statistics Denmark. Additionally the concept of 'relatives' in Danish ('pårørende') is a fuzzy concept (cf. Lakoff 1973) that means different things to different people and is highly dependent of the context in which it is used. Furthermore, the topic of caregiving for a relative in need of care is highly emotional and requires delicate communication – both when engaging with stakeholders and in the communication surrounding the final outcome.

In this paper we highlight some of the encountered challenges and point to a few concrete outcomes that may be useful for other producers of brand-new social statistics, particularly statistical producers who intend to involve stakeholders in their development process.

Keywords: New Statistics; Relatives; GSBPM; External involvement; Fuzzy concept

1. Introduction

During a 3-year period (2022–2025) Statistics Denmark took in the task of developing a *register on relatives* with the purpose of aiding the civil society and research community in conducting analyses of relatives (*‘pårørende’* in Danish; in the remainder of this paper we will primarily use the Danish term. The reason for this follows) broadly speaking and the caregiving resources that these relatives potentially provide more specifically.

The main result is the register on relatives which is updated yearly, and from which microdata is available for researchers and others through Denmark’s Data Portal just as users can purchase customised solutions. For the general public use, Statistics Denmark also publish a range of tables in the online StatBank. Some of these are population tables counting who has, and thus also who has no, relations in the Danish population. Several more specialised tables are also published, such as the propensity to visit with one’s doctor split on the type of relatives one has.

The project was founded by a non-profit philanthropic foundation, TrygFonden, and the development of the new register was based on two overall principles: one that it should be founded on available, administrative register data, and two, that the construction of the register was to be carried out in cooperation with the coming users of the statistics.

The interaction between these two principles posed several benefits and challenges during the development project; particularly since the core population of *pårørende* is a rather fuzzy concept that is not easily identified through administrative data. In this paper, we will outline some of the encountered challenges and point to some key learning points that may be useful for other producers of brand-new statistics.

2. What is a fuzzy concept?

The term ‘fuzzy concept’ stems from the linguistic sciences and its semantics was originally coined by the American linguist and philosopher George Lakoff in 1973. Numerous other linguists and scientists have since expanded on the conceptualization but it remains that a fuzzy concept refers to a concept that is difficult to define and is context-dependent. The opposite is a ‘crisp concept’ which in contrast has rigid boundaries and thus is uncomplicated to define. When developing statistics and

perhaps especially statistics based on register-data, the latter is optimal but as producers of statistics can testament, even the apparent most straightforward concept can be difficult to define in exact terms. The concept of a child is one such example. While a child is often an individual below the legal age of 18, which can be defined in binary terms, it becomes more difficult to define when considering the children in a family-context. Does this only include children under the age of 18 (i.e. a binary boundary); should older children who still live in the home of their parents be included (and until which age; another binary boundary); and how to treat children in a spilt family context? Therefore the contrast of fuzzy concepts versus crisp concepts quickly becomes blurry in praxis.

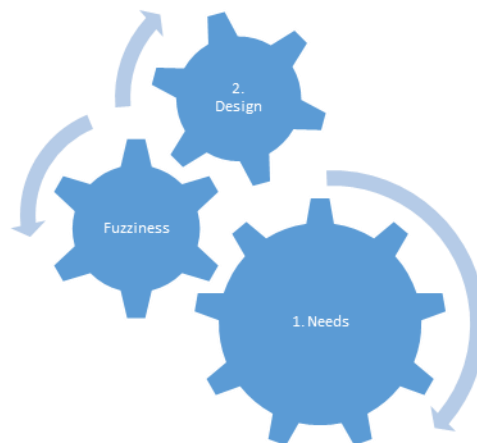
However, the concept of relatives which is called *pårørende* in Danish, holds other levels of fuzziness. In Danish, the term *pårørende* is a much broader term than the level of kinship that is implied in both our chosen English translation used for the statistics ('relatives') and other translations that were considered (next-of-kin; family carers). Where the English term 'relative' implies a kinship relation, the Danish term doesn't necessary do so. In Danish a *pårørende* is someone who has a close relation to the main person, but the term itself does not imply that it has to be a family relation. It could be a close friend or a neighbor that takes on the role of being a *pårørende* for the main person. However, we also observe a shift in the Danish language so the term *pårørende* is increasingly used when referring to non-specified family members in general – which obviously adds to the linguistic fuzziness surrounding the development of this new statistics.

In addition to this linguistic context that surrounded the development of the new *register on relatives*, a *pårørende* in the sense of the developed statistics also contains the understanding that the *pårørende* takes on a task or a role in relation to the main person. To illustrate this, a person may be a main person's daughter; i.e. understood as a precise family relation, but in a situation where the main person becomes hospitalized, the daughter takes on the role of being a *pårørende* to him/her. Additionally, there is also a temporal aspect here in that the shift from considering oneself as a daughter to rather being a *pårørende* to the main person does not happen immediately but is rather a role one grows into during the course of e.g. a progressing illness and the increased need for support.

3. A fuzzy concept in a GSBPM framework

When tasked with making statistics on a fuzzy concept, challenges and opportunities arise. Working within the GSBPM approach constitute an opportunity as it provides a framework for consulting user needs, identifying concepts from a user perspective, and checking data availability. The challenge is to design a definition that holds enough fuzziness to be accepted by the users while being concrete enough to be put into a statistical framework. In figure 1 we depict the way we experienced the interaction between fuzziness, the needs phase, and the design phase of the GSBPM as linking into each other as gears.

Figure 1: Fuzziness in a GSBPM framework



In the project we spent the better part of the first project year identifying and clarifying needs to get a grasp on the fuzziness and also handle user expectations for the final result. Because we prioritized taking the time needed working through the steps of the needs phase, it is our postulate that getting through the design phase went smoother. However, when you start designing the statistical system, you might need to reconsolidate user-needs, especially if you have to make changes or assumptions that do not entirely match the user-needs. In the following sub-chapters, we will dive into some key take-aways.

a. Take-away: Identifying user needs and making a communication strategy

Before consulting user-needs, we conducted a literature review on the topic of *pårørende* in order to identify a potential definition of *pårørende*. This review did not yield a clear or operational definition of *pårørende*. Consequently, we identified a broad group of relevant stakeholders—many without experience in statistics or register-based data systems—and invited them to participate in a focus group. The aim was, amongst others, to identify who *pårørende* typically are from a holistic perspective. A central methodological challenge was then how to translate qualitative, real-world knowledge into inputs suitable for designing register-based statistics upon.

The discussions revealed perspectives that would likely not have emerged, had we only consulted traditional users of Statistics Denmark. For example, one NGO noted that among its target group, *pårørende* are often non-family members, while others highlighted the role of children-in-law. Nevertheless, there was broad consensus among a broad stakeholder group that *pårørende* are most commonly close family members.

As a result, we also reframed the statistics conceptually as a register and statistics of *potential pårørende*, recognizing that the register would not identify who the actual *pårørende* in specific circumstances are. It would also not be possible to identify the *pårørende* who are not a family relation in a register data context.

From this point onwards, we used this framing as a communication strategy both during the project and in the communication to future users, signaling that we are aware that there are some parts of the broader understanding of *pårørende* that is not possible to handle in a register and statistics based on the current available administrative data-sources.

b. Take-away: Identifying concepts

When moving on to identifying concepts, we formed a working group consisting of expert users of Statistics Denmark products. Discussions within this working group focused on register content and structure. The participants supported a narrow definition of *potential pårørende* including only close family members.

When asked how one could concretize the fuzzy concept of *pårørende*, expert-users pointed to geographical distance as a key indicator of the ability to act as *pårørende*. Co-residence, current or historical, was also identified as an important variable.

If we could succeed in implementing these concepts in the register, it would embed some of the fuzzyness and thus meet userneeds for the future application of the register content.

c. Take-away: Checking data availability

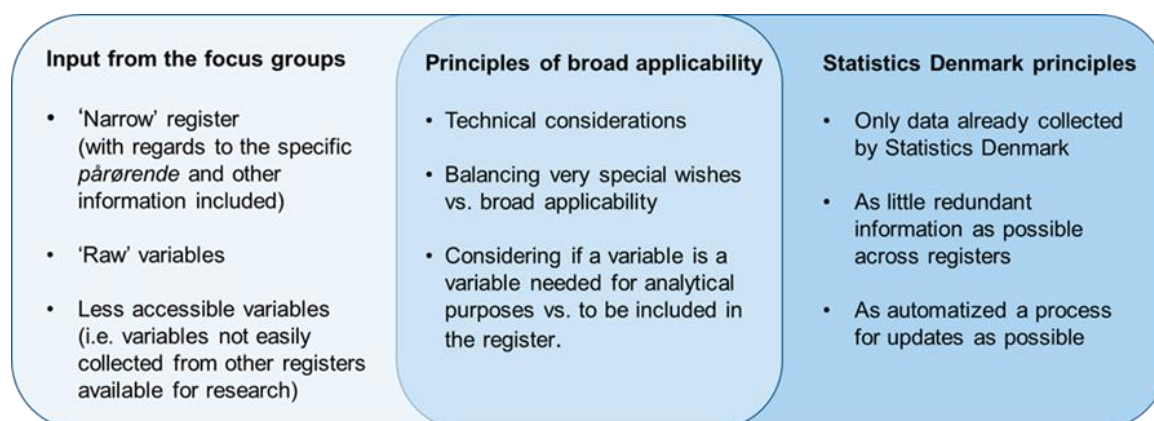
The group of expert users also pointed out a number of potential supplementary data sources. These sources included health sector registrations, powers of attorney, network analyses, and multigenerational linkage data. However, none were sufficiently mature or accessible to be included within the project timeframe.

Although it was not in the scope of the project to collect new data sources, it did however underline the need for a flexible register design capable of incorporating future data sources.

d. Take-away: Design phase

Based on the needs phase, the project team formulated a set of design guidelines, which were approved by the steering committee. These guidelines balanced stakeholder input, the requirement for a broadly applicable yet conceptually narrow register, and Statistics Denmark's principles for data ethics, see figure 2.

Figure 2: Principles from needs phase to design phase



4. Fuzziness also means changes in concept over time

When working with a fuzzy concept nothing is set in stone, and you must create a flexible statistical system. This was emphasized as we experienced a shift in the public understanding or use of *pårørende* during the project years. The public and institutional understanding of *pårørende* evolved. The concept shifted from viewing *pårørende* primarily as a resource, to recognizing their physical and psychological burden, and finally towards a broader and more traditional understanding of *pårørende* as family members, i.e. relatives.

We also experienced an increased focus on the population without *pårørende*. Our system was already designed as a population register holding the entire population, and hence also included those with no *pårørende* in the Danish population.

Had we not designed a flexible system these changes might have posed a real problem for the adaptability of the statistics.

5. Conclusion

In the development of the new *register on relatives*, the GSBPM model has proven to be a useful framework for the involvement of external stakeholders in the development of new statistics. Hopefully, we have also conveyed the message that it is important to continuously reconsider previous GSBPM step in the subsequent design phase. Furthermore, it requires careful communication to involve the future users in the development of a new register, especially when addressing a sensitive and fuzzy concept such as *pårørende*.

6. Appendix: Further reading

Lakoff, George (1973): *Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts*, Journal of Philosophical Logic, Vol. 2, pp. 458-508

Opinion / Helsedirektoratet (2021): *Nasjonal pårørendeundersøkelse*,
<https://l1nq.com/pfvjq7n>

Statistics Denmark (2025): *Documentation of the Register of Relatives and the associated Relatives Statistics*,
<https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/relatives>

Statistics Denmark (2026): *Statistics on Relatives in the StatBank*,
<https://www.statbank.dk/21200>

Statistics Denmark (2026): *Subject page for Statistics on Relatives*,
<https://www.dst.dk/en/Statistik/emner/borgere/husstande-og-familieforhold/paarorende>

VIVE (2024): *At være pårørende i Danmark – en kvalitativ analyse af pårørendes erfaringer fra et hverdagslivsperspektiv*, <https://www.vive.dk/media/pure/rv74poov/24963102>