Final report

Improving Equality Data Collection in Belgium II

UNIA
Together for equality

Equal Opportunities.be

Table of contents

We would like to express our heartfelt thanks to the individuals and organisations who participated throughout the research process leading to this report. It is thanks to their time, feedback, and knowledge that we have been able to realise this work. This collaboration and these discussions have been extremely valuable.

1. Introduction

This report is the result of the project *Improving Equality Data Collection in Belgium II* (IEDCB II). This research project was carried out by Unia between January 2023 and June 2024 and was funded by the Equal Opportunities Team of the Federal Public Service (FPS) Justice. It is a follow-up project that builds on the results and recommendations of the IEDCB I project (2021)[1].

IEDCB II focused on improving the collection of *equality data* for the following grounds of discrimination: disability, health status, and physical characteristics. In this introduction we first provide more information about equality data and outline the context and objectives of the project. The subsequent chapters describe the methodology, results, and recommendations.

1.1. What are equality data?

The term 'equality data' refers to **any data that are useful for the purposes of describing and analysing the state of equality or inequality in society**[2].

'Data' is used as a general term and can refer to both quantitative and qualitative data. In some cases, it refers to data collected for the specific purpose of measuring (in)equality or discrimination. However, it may also refer to data that were collected for other purposes but that can also be used to measure inequality, for example because they allow us to compare the situation of different groups.

More specifically, the following types of data constitute important sources of equality data[3]:

- Data collected through a population **census**
- Data collected through **surveys**; these can be official, large-scale surveys (e.g., the Labour Force Survey), but also smaller, more targeted surveys
- Data from **administrative databases** (e.g., the Crossroads Bank for Social Security)
- Results of **discrimination testing** (e.g., field experiments conducted in the housing market)
- **Complaints data** (e.g., reports received by Unia)
- Outcomes of **diversity monitoring** (e.g., diversity of the workforce in the public sector)
- Results of **qualitative research** (e.g., in-depth interviews, focus groups)

Why are equality data important? **Accurate and reliable equality data are indispensable in the fight against discrimination and inequality**. In order to develop policies based on facts rather than on intuition, it is important to have an accurate picture of reality. Equality data can be used for the following purposes, among others[4]:

- **Monitoring**: in which domains of society do inequality and discrimination occur, who is affected (see box 'Grounds and domains'), and how does the situation evolve?
- **Making visible**: some methods (e.g., field experiments) can reveal forms of discrimination that are hidden or not often reported
- **Understanding**: understanding the structural mechanisms that lead to inequality and discrimination

[1] Unia (2021). Final report 'Improving equality data collection in Belgium'.
[2] European Commission (2016). European Handbook on Equality Data - 2016 Revision, p. 15.
[3] For more information on these types of data, see European Commission (2016). European Handbook on Equality Data - 2016 Revision, p. 31-36.
[4] Subgroup on Equality Data (2018). Guidelines on improving the collection and use of equality data.

- **Developing and evaluating legislation and policies**: based on the above information, evidence-based policies can be developed and their impact can be evaluated
- **Raising awareness**: data can help to raise awareness regarding the topic of non-discrimination
- **Providing evidence**: equality data that demonstrate discrimination can be included as evidence in legal cases concerning discrimination
- **Reporting**: assessing to what extent Belgium complies with the obligations outlined by international human rights treaties (such as, for example, the *United Nations (UN) Convention on the Rights of Persons with Disabilities*, which we discuss further in section 1.3.1)

In Belgium, quite a lot of data are collected that can serve as equality data. However, these data are scattered across numerous bodies and organisations, which is partly due to the complex distribution of jurisdiction between the different governments in Belgium. Moreover, these various initiatives to collect equality data are not coordinated centrally (i.e., at the national level). As a result:

- many different definitions and methods are used
- data on certain groups or domains are lacking
- it is difficult to know which data exist and where to find them, resulting in insufficient use of the available data.

This **fragmentation of Belgian equality data** was the starting point for the IEDCB project.

Grounds and domains

Throughout this report, we talk about 'grounds' and 'domains'. What do we mean by these?

What are grounds? Anti-discrimination legislation (federal laws, decrees, and orders) talks about 'protected criteria', also called discrimination grounds[5]. Discrimination based on these grounds is forbidden and punishable. They include, for example, skin colour, sexual orientation, disability, etc. In this report, we focus on three discrimination grounds: disability, health status, and physical characteristics (see section 1.3 for definitions). We will refer to these grounds as the 'current grounds'.

What are domains? Discrimination and inequality can occur in different domains of life. One can be discriminated against at work (e.g., when looking for a job), in education (e.g., in the classroom), in the context of housing (e.g., when trying to rent a house), etc.

It is relevant to know *who* is discriminated against (on the basis of which ground) as well as *where* that discrimination occurs (in which domain). That is why we often look at equality data through the lens of grounds and domains. For example, a source may contain data on inequality or discrimination based on the ground of 'disability' in the domain of 'employment'.

1.2. Context and objectives of the project

The first IEDCB project, which we refer to in this report as IEDCB I, ran from 2020 to 2021. Both this first project and the current project (IEDCB II) drew heavily on work carried out at the EU level[6], in particular by the *Subgroup on Equality Data* and by the *Fundamental Rights Agency* (FRA), which

---

[5] Unia. Grounds of discrimination.

[6] For an overview of the work on equality data at the EU level, see the European Commission website.

facilitates the work of the subgroup. Throughout both projects, the *European Handbook on Equality Data* and the *Guidelines on Improving the Collection and Use of Equality Data* served as central guides[7].

The first project (IEDCB I) was restricted to equality data for three groups of discrimination grounds: (1) the so-called 'racial' grounds: so-called 'race', skin colour, nationality, ancestry, and national or ethnic origin; (2) philosophical or religious beliefs; (3) sexual orientation, gender identity, gender expression, and sex characteristics[8]. This was done in order to keep the project feasible, with the intention of later extending the work to include other grounds. This extension was therefore one of the main objectives of the current follow-up project (IEDCB II).

The IEDCB II project built on IEDCB I. Specifically, the **following results** were achieved:

1. **An extension of the inventory of equality data in order to include the grounds of disability, health status, and physical characteristics**. During IEDCB I, existing sources of equality data were gathered and mapped for a limited set of grounds. During IEDCB II, this inventory was expanded to also include equality data for the three current grounds.
2. **An update of the existing inventory**. As the IEDCB I project ended in 2021, an update of the existing inventory was carried out during IEDCB II. This involved reviewing the sources (including checking whether they contained any data on the current grounds), replacing them with more recent versions whenever necessary, and adding new sources that had been published since the previous project.
3. **A renovation of the *data hub*.** During IEDCB I, a publicly accessible website was created based on the inventory, which allows users to search for the equality data they need by using various filters (see box 'The Equality Data Hub'). During IEDCB II, the new and updated sources were added to the data hub. Some new features were also added, such as the possibility to filter on sources containing intersectional data or analyses.
4. **An overview of gaps and recommendations.** Based on the new inventory and on an extensive consultation of various stakeholders, important gaps and needs were identified for the existing equality data covering the grounds of disability, health status, and physical characteristics. These results are described in this report and then linked to recommendations for the future.
5. **A study on self-identification of origin.** The methodology and results of this study, which focused on equality data for the so-called 'racial' grounds (so-called 'race', skin colour, nationality, ancestry, and national or ethnic origin), are described in a separate report[9].
6. **An initial reflection regarding a platform for equality data.** One of the recommendations of IEDCB I was to establish a consultative body with relevant actors in order to develop a coordinated and structural policy for equality data. During IEDCB II, an initial reflection was conducted on how such a platform could be established (by the Equal Opportunities Team of the FPS Justice).

In the 'Methodology' section, we describe how results 1 to 4 were realised.

[7] European Commission (2016). European Handbook on Equality Data - 2016 Revision.
Subgroup on Equality Data (2018). Guidelines on improving the collection and use of equality data.
[8] For more information on these grounds and their definitions, please refer to the final report of IEDCB I: Unia (2021). Final report 'Improving equality data collection in Belgium'.
[9] Unia (2024). Study on self-identification of origin in Belgium.

EQUALITY DATA HUB

THE EQUALITY DATA HUB: AN EASY WAY TO FIND DATA ON (IN)EQUALITY AND DISCRIMINATION!

Looking for data on discrimination and (in)equality in Belgium? Then be sure to take a look at the Equality Data Hub: https://www.equalitydata.unia.be/fr/ (available in French or Dutch).

Who is the data hub for?

The aim is to make it easier for everyone to find data on discrimination and inequality. However, the tool is specifically aimed at policy makers, journalists, researchers, students, as well as anyone working on the topics of discrimination, inequality, diversity, and inclusion.

How do you use the data hub?

You can search the database for existing data on discrimination or (in)equality based on a number of protected discrimination grounds. You can use filters to look for sources covering **specific grounds** and **specific domains**.

An example: You need data on inequality based on disability on the labour market. Then you can select 'Disability' for the 'Grounds' filter and 'Work' for the 'Domains' filter. This will provide you with a list of all relevant sources.

There are **plenty of other filters** you can use to further refine your search. For example, you can filter by the region of Belgium which the source covers or search for sources that contain intersectional data or analyses. You can also search by keywords.

1.3. Selection of discrimination grounds and their definitions

In this project, we focus on three grounds of discrimination: disability, health status, and physical characteristics. Why did we select these grounds and which definitions did we use?

1.3.1. Disability

We define **disability** in the same way as in the **UN Convention on the Rights of Persons with Disabilities**: "Persons with disabilities include those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others."[10]. We therefore understand disability as a **long-term** situation that arises from the **interaction** between a personal impairment and an environment that is insufficiently adapted.

Extending the work on equality data to the ground of disability is important for several reasons.

[10] United Nations (2006). Convention on the Rights of Persons with Disabilities, art. 1.

First, the 2014 **review of Belgium with regard to the UN Convention** concluded that Belgium does not collect data on the situation of persons with disabilities in a sufficiently systematic way[11]. Belgium ratified this convention in 2009 and thus committed itself to promoting the rights it lists. In 2011, Unia was designated as an independent mechanism responsible for monitoring the implementation of this convention[12]. To assess Belgium's compliance with the Convention, data that shed light on the situation of persons with disabilities in Belgium (i.e., equality data) are crucial[13]. In Article 31 (Statistics and data collection), the Convention therefore explicitly calls on states parties to "**collect appropriate information, including statistical and research data, to enable them to formulate and implement policies to give effect to the present Convention**"[14]. Such data should allow to monitor the impact of both disability-specific policies and the impact of mainstream policies on persons with disabilities (the so-called 'twin-track approach')[15]. In summary, Belgium is obligated to collect, make available, and use equality data on disability to improve the situation of persons with disabilities. However, even today the situation in terms of data collection still leaves much to be desired, as Unia and civil society have pointed out during the run-up to the next evaluation[16].

The Federal Action Plan Disability (2021-2024)[17] contains some measures to improve this data collection, including the current IEDCB project (measure 129). A **Working Group on Statistics** was also created in 2023 within the framework of the Interministerial Conference (IMC) Disability (measure 130 in the action plan), with the aim of strengthening the statistical capacity related to persons with disabilities[18]. This working group (in which Unia participates) works mainly on large-scale data sources and does not specifically focus on inequality and discrimination, but still identifies important gaps in existing (equality) data. Where relevant, this report will refer to the results of the working group.

In addition, Unia receives many **reports** related to the ground of disability[19]. In 2022, there were 1,059 reports of potential discrimination, representing almost 20% of all reports under Unia's jurisdiction. If we look at the **case files** that were actually opened, the percentage came up to 26% (543 case files).

1.3.2. Health status

The second ground we included is **health status**. By this we mean the current health status (e.g., a physical or mental condition), the past health status or disease history (e.g., whether one has ever had cancer), or the future or reasonably foreseeable health status (e.g., a degenerative disease that will increase in severity in the future)[20]. Our definition does not require the condition to be linked to an

---

[11] Committee on the Rights of Persons with Disabilities (2014). Concluding observations on the initial report of Belgium.
[12] For Flemish competences, this role has been with the Flemish Human Rights Institute since 15 March 2023.
[13] European Union Agency for Fundamental Rights (2023). Implementing the UN Convention on the Rights of Persons with Disabilities: Human rights indicators.
[14] United Nations (2006). Convention on the Rights of Persons with Disabilities, art. 31.
[15] Committee on the Rights of Persons with Disabilities (2016). Rules of procedure, Annex: Guidelines on independent monitoring frameworks and their participation in the work of the Committee on the Rights of Persons with Disabilities, p. 33-43.
European Union Agency for Fundamental Rights (2023). Implementing the UN Convention on the Rights of Persons with Disabilities: Human rights indicators.
[16] Unia (2024). Parallel report to the Committee on the rights of persons with disabilities, p. 34-35.
Belgian Disability Forum (2018). Alternative report to the Committee on the Rights of Persons with Disabilities within the framework of the 2nd and 3rd review of Belgium, p. 38-40.
[17] SPF Sécurité sociale (2021). Plan d'action fédéral Handicap (2021 -2024).
[18] SPF Sécurité sociale (2023). Groupe de travail statistiques - CIM Handicap - Note finale.
[19] Unia (2022). Annual report: section 'Unia, a body promoting equality'.
Unia. Chiffres d'Unia.
[20] Unia. Grounds of discrimination.

official diagnosis. This is a deliberate choice based on our stakeholder consultation and the fact that equality data are often based on self-reported health status.

The **distinction between disability and health status** is not always clear-cut. The distinction can be made on the basis of the duration and the situational character: if the condition is **long-term** (often a threshold of six months is used[21]), and if there are **participation problems** due to its interaction with an unadapted environment, it actually falls under the definition of disability. This will often be the case for chronic diseases or mental illnesses, which can have a major impact on the person's participation in society (e.g., Alzheimer's disease, haemophilia, etc.) By health status, we thus mainly refer to situations that are limited in time (e.g., COVID-19, a broken leg, etc.) and/or to situations where there are no participation problems but one is discriminated against on the basis of a condition (e.g., AIDS, hepatitis C, etc.).

For the criterion of **health status**, Unia also received numerous reports in 2022 (1326) and opened many case files (181). Although these figures are partly related to the temporary measures taken during the COVID-19 pandemic, discrimination on the basis of health status is also a significant problem outside of these exceptional circumstances. It is unclear, however, what other sources of equality data exist on this topic. Moreover, the distinction from the ground of disability is not always easy to make and data on these two grounds often appear in the same sources. It is therefore relevant to look at equality data on health status and disability within the same project.

1.3.3. Physical characteristics

Finally, we also include the criterion of **physical characteristics**. For this we use the following definition: any external feature that is potentially stigmatising for the person (and that is not linked to other discrimination grounds such as, for example, skin colour, religion, disability, or health status). This could include a birthmark, a scar, the shape or size of the body, ... Based on the stakeholders' input, we decided to use a broader definition for this particular project than the strictly legal definition that is also used by Unia[22]. Indeed, in our search for relevant data, we did not *a priori* limit ourselves to discrimination based on characteristics that occur outside a person's will. Unia regularly receives reports about 'voluntarily chosen' physical characteristics (e.g., tattoos or piercings) and research shows that these are indeed linked to stigmas[23], so we did not want to automatically exclude any available data on this from the inventory.

Terms such as *fat phobia*[24]*, face equality*[25]*,* and *lookism*[26] suggest that discrimination based on physical characteristics is a real problem. Some case files of Unia also show clear discrimination based on this ground (e.g., a recent case in which a woman was fired because of a dental problem[27]). Overall, however, Unia receives relatively few reports related to physical characteristics and it is unclear how often discrimination based on this ground occurs.

[21] Eurostat. Glossary:Disability.
[22] Unia. Grounds of discrimination.
[23] Baert, S., Herregods, J., Sterkens, P. (2024). What does job applicants' body art signal to employers? *Journal of Economic Behavior & Organization, 217*, p. 742-755.
[24] See, for example, Puhl, R. (2021). Weight stigma is a burden around the world – and has negative consequences everywhere. *The Conversation.*
[25] See, for example, the website of Face Equality International.
[26] See, for example, Minerva, F. (2019). Onzichtbare vorm van discriminatie: onaantrekkelijke mensen moeten opboksen tegen diepgewortelde vooroordelen. *Knack*.
[27] Unia (2023). Une dame licenciée pour un problème dentaire gagne en justice.

Again, data sources do not always make a clear **distinction** with disability or health status. For example, there are forms of discrimination that could possibly legally fall under the ground of disability (e.g., discrimination based on weight), but that are sometimes covered in research under the heading of physical characteristics. Given the great lack of knowledge about discrimination based on this ground and the possible overlap with data on disability and health status, we decided to include physical characteristics in this project as well.

1.4. Handling data in a responsible manner: legislation and general principles

Equality data are in many cases about individual persons. Therefore, unless the data are completely anonymous and cannot in any way be traced back to any identifiable person, they are **personal data**. Even when it comes to completely anonymous data, the data that need to be collected in order to compile equality data are often of a **sensitive nature** (e.g., data on health, origin, sexual orientation, etc.) and there are important ethical considerations. We would therefore like to briefly reflect on some elements that are crucial in the context of equality data.

First, we highlight the **legal framework applicable to the processing of personal data**. The final report of IEDCB I contains a clear overview of the relevant international and Belgian laws and how they relate to each other[28]. We refer to this report for more details and only reiterate the main conclusions here:

- There is a **clear legislative framework**, which provides important guarantees for the protection of personal data in data collection and processing.
- In general, it is not allowed to collect **'special categories' of data**[29]. However, this is allowed in certain situations, for example if the person has given explicit consent or if the processing is necessary for the public interest, scientific/historical research, or statistical purposes[30].
- A special case concerns **Statbel**, the national statistics office, which is prohibited by the Belgian law of 4 July 1962 on public statistics[31] from publishing statistics concerning the private sphere (which largely correspond to the special categories of data). This provision (which is indeed more strict than international law) is often interpreted in a very restrictive sense, even by bodies or organisations that are not covered by this law.
- In sum, there are certainly opportunities to collect **equality data** within this legislative framework, as long as the general regulations are taken into account and the data are collected for a **clear and legitimate purpose**. Transparency and communication about this is crucial when collecting equality data.

In addition, human rights in general should always be respected when collecting and disseminating (equality) data. For this, we refer to the **human rights-based approach to data** outlined by the Office of the High Commissioner for Human Rights (OCHCR)[32]. This framework was also discussed in more detail in IEDCB I, but we would like to reiterate its six central principles here:

---

[28] Unia (2021). Final report 'Improving equality data collection in Belgium', Chapter 3: The legal framework: data protection, p. 16-17.
[29] "revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, (…) data concerning health or data concerning a natural person's sex life or sexual orientation"
[30] Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Official Journal of the European Union 04 May 2016, L119/1, art. 9.
[31] Loi du 4 juillet 1962 relative à la statistique publique.
[32] Office of the High Commissioner for Human Rights (2018). A Human Rights Based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development: Guidance Note to Data Collection and Disaggregation.

- **Participation**: The groups that are the subject of the data should be involved throughout the process of data collection and processing.
- **Data disaggregation**: The data should be collected in such a way that they can be disaggregated to compare different groups, in order to measure inequality or discrimination (e.g., the data should contain a variable that can distinguish between persons with and without disabilities).
- **Self-identification**: Where possible, data on personal characteristics should be provided by the individuals themselves, giving them maximum choice on whether and how they wish to do so. We discuss self-identification in more detail in the Operationalisation chapter.
- **Transparency**: The way the data are collected and processed should be transparent. Moreover, the data should be disseminated in an accessible way.
- **Privacy**: Data should be collected and processed in a secure manner, protecting the privacy of the individuals involved and preventing their identification.
- **Accountability**: Data should be collected responsibly (i.e., with respect for human rights) and used to hold government and other actors accountable on human rights obligations.

Finally, **the general principle of *'do no harm'*** always applies: equality data should only be used for the benefit of the groups concerned and of society as a whole, and should therefore never be used to discriminate against or exclude these groups[33].

The above principles should always guide the collection of equality data and will be referred to regularly throughout this report.

1.5. Content of this report

The next chapters will cover the following topics:

- In **Chapter 2**, we describe the **methodology** we used for this project. We explain how we realised the inventory of data sources and how we gathered information from key stakeholders.
- In **Chapter 3**, we discuss the results of this process: which **gaps and needs** can we identify for equality data on these three discrimination grounds (disability, health status, and physical characteristics)? For this purpose, we analyse the inventory and summarise the gaps and needs brought up by the stakeholders.
- In **Chapter 4**, we look at how the discrimination grounds are **operationalised** in Belgium and outline different methods and perspectives in this regard. In the first part of this chapter, we give an overview of which methods are used in Belgium to measure disability and discuss the advantages and disadvantages of each of these methods. In the second part, we address the confounding of terms related to weight-based discrimination and the definition of 'obesity'.
- In **Chapter 5**, we formulate our **recommendations**. Here, we focus mainly on policymakers, producers of equality data, and potential users of the data.

We realise that the report is written quite concisely. Unfortunately, we cannot cover all topics in detail and sometimes have to limit ourselves to the issues that came up most frequently in the stakeholder consultation. Moreover, there are certainly actors whom we have not been able to consult and therefore we may have overlooked certain issues. However, our aim with this report is to capture the main points and to set out clear paths for further work.

[33] Office of the High Commissioner for Human Rights (2018). A Human Rights Based Approach to Data - Leaving No One Behind in the 2030 Agenda for Sustainable Development: Guidance Note to Data Collection and Disaggregation.

2. Methodology

In this chapter, we present the methodology used for this project. First, we describe the methodology that we followed for **mapping** the existing equality data for the current grounds. Then, we explain in more detail the **crowdsourcing exercise** that we organised to collect information related to the three current discrimination grounds, through a consultation of different actors working on these topics.

2.1. Inventory of existing sources

We created an **inventory** of existing sources of equality data for Belgium related to the grounds of disability, health status, or physical characteristics. For this we used the mapping tool[34], which was developed by the *Subgroup on Equality Data* and adapted to the Belgian context during the IEDCB I project. The inventory also forms the content of the data hub.

2.1.1. Search and selection of sources

Since the definition of equality data is very broad (see section 1.1), a very large number of sources can be considered to contain equality data for the current grounds. However, it is necessary to be able to get a coherent picture of these sources that allows to analyse equality or inequality for a given discrimination ground. Therefore, we reflected on how to arrive at a coherent and practical inventory. Below, we describe in detail how we selected sources for the inventory. We determined **specific selection criteria**, always considering the following question (with the data hub in mind): "*Does this source demonstrate equality or inequality for any of the grounds?*" We also restricted ourselves to sources containing data on Belgium or on a part of Belgium.

We gathered **potential sources** in several ways. First, we widely distributed a survey to people working on the current discrimination grounds and/or working on data collection (the survey was part of the crowdsourcing and is explained in section 2.3.2). The survey already yielded around 100 potential sources and provided a starting point for us to proactively search for additional sources ourselves. In order to do this thoroughly, a list of websites of potential equality data producers was compiled. The sources we identified by searching these websites were then analysed more closely to check whether they met our selection criteria. Finally, the selected sources were included in the inventory.

To decide whether or not to include a source in our inventory, we adhered to the following list of selection criteria, which is largely similar to the one used during IEDCB I.

1) Does the source cover at least one of the grounds?

The data source had to cover **at least one of the following grounds of discrimination:** disability, health status, or physical characteristics. Because we were also updating the inventory created during IEDCB I, we additionally included data sources on the grounds covered in IEDCB I: so-called 'race', skin colour, nationality, ancestry, national or ethnic origin, philosophical or religious beliefs, sexual orientation, gender identity, gender expression, and sex characteristics.

2) Does the source contain any of the following types of data?

The source had to contain one of the following **types of data**:

- Data collected through a population census
- Data collected through (large-scale or small-scale) surveys

[34] We explain the 'mapping tool' in more detail in section 2.1.2.

- Data from administrative databases
- Results of discrimination testing
- Complaints data
- Outcomes of diversity monitoring
- Results of qualitative research

Sometimes a source contained quantitative data that did not fall into any of the above categories but that nevertheless seemed particularly relevant to us, such as data on the number of government websites that meet digital accessibility standards and are therefore accessible to persons with disabilities. In such cases, we still included the source and registered it as 'other quantitative data' in our inventory.

For sources containing qualitative data, we applied an additional condition: we only included sources that were based on clear social science research methods, such as in-depth interviews, focus groups, and so on. Therefore, we did not include sources that took the form of legal or policy analysis.

3) Was the source published in or after 2010?

To provide the most up-to-date data possible in the data hub, we decided to restrict ourselves to sources that were published **in or after 2010**.

4) Is the source freely and publicly available?

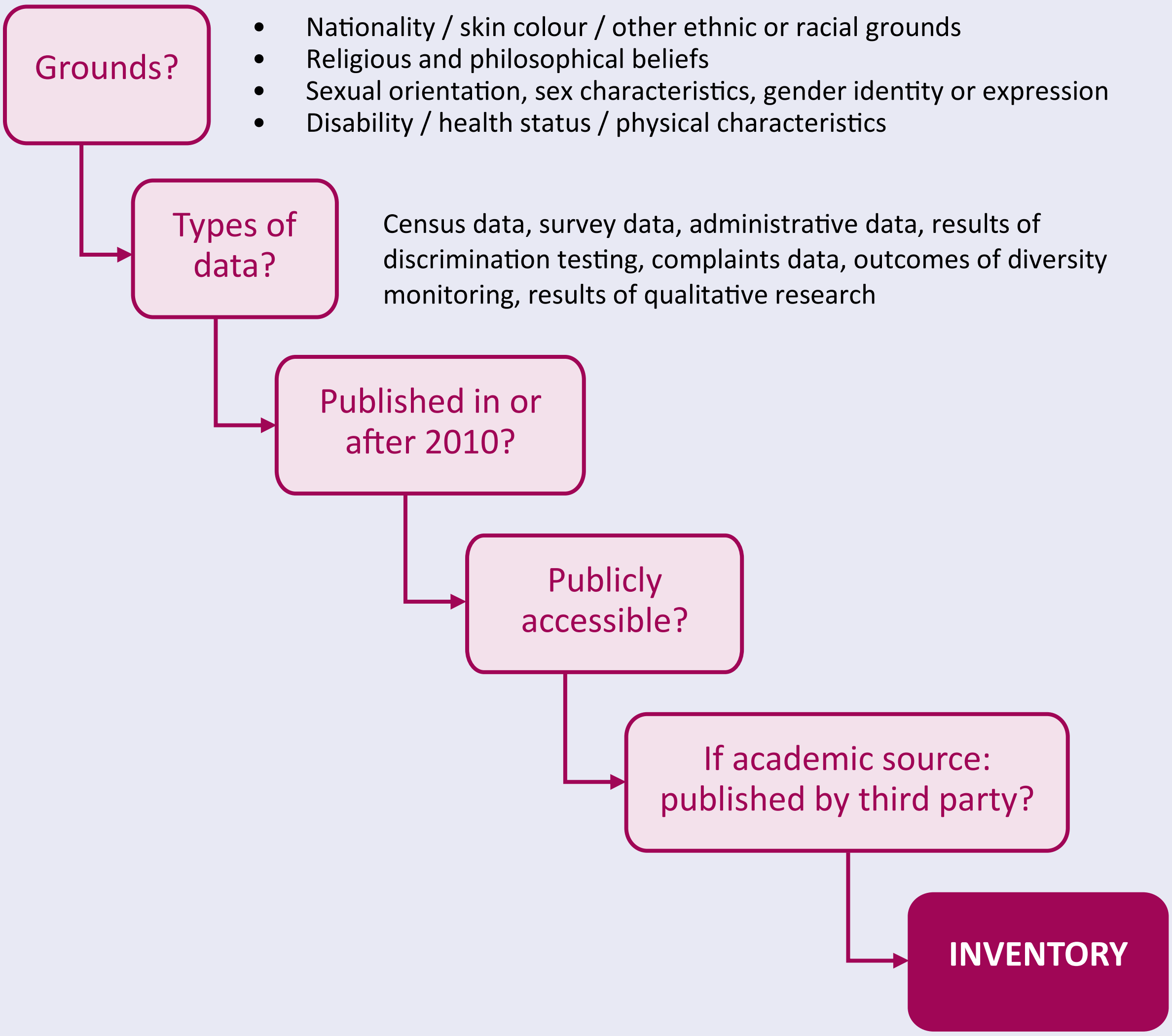
The purpose of the data hub is that anyone can access the included equality data. When using the data hub, the user is referred to each source via a link to the website of the producer/publisher. We therefore decided to include only sources that are **publicly accessible**. If one can access the source through a simple request on the website, we also included the source.

5) In case of an academic source: is the source published by a third party?

Unfortunately, it was impossible to include all results of **academic research**, due to the large quantity and due to them not always being publicly accessible. We therefore decided to include only academic research that was explicitly funded, commissioned, or published by a third party (e.g., government administrations or civil society).

These different criteria are summarised in the decision tree depicted in Figure 1.

FIGURE 1. DECISION TREE FOR SELECTING SOURCES OF EQUALITY DATA FOR THE INVENTORY.



2.1.2. Working method for the inventory

The selected sources then had to be inventoried. We did this by using the *mapping tool*, a tool that was developed by the *Subgroup on Equality Data* and that allows to create an inventory of equality data[35]. It takes the form of an Excel file intended to facilitate the inventory and categorisation of sources of equality data. During the IEDCB I project, the original tool of the *Subgroup on Equality Data* was adapted to the Belgian context. During the current project (IEDCB II), we further adapted this working tool. Table 1 provides an overview of the information we recorded for each data source.

[35] Subgroup on Equality Data (2019). Compendium of Practices on Equality Data: Equality Data Diagnostic Mapping Tool.

TABLE 1. CATEGORIES OF INFORMATION RECORDED DURING THE INVENTORY.

Source characteristics	Name of the source
	Languages in which the source is available
	Author
	Year of publication
	Link to the source (or links if the source exists in multiple languages)
Characteristics of the data	Type of data
	Geographical coverage
Discrimination grounds	Nationality
	Skin colour
	Origin/migration background/ancestry/other so-called 'racial' grounds
	Religious and philosophical beliefs
	Sexual orientation
	Gender identity
	Gender expression
	Sex characteristics
	Disability
	Health status
	Physical characteristics
Domains[36]	Work
	Education
	Health and well-being
	Poverty and standard of living
	Housing and living situation
	Mobility and transportation
	Justice and police
	Media
	Digitalisation
	Other domain (accompanied by a description of these other domains)
	Purely demographic data (no domain)
Specific aspects	Does the source contain data on experiences of discrimination?
	Does the source contain data on hate speech or hate crime?
	Does the source contain intersectional data or analyses? (accompanied by a description of these data/analyses)
Keywords	Maximum of 10 keywords

[36] This refers to the domain of life that is covered by the data, such as work, education, media, etc. The domains that we record evolved between the first and second phases of the IEDCB project. Indeed, based on discussions with the advisory group, we decided to add some domains that seemed particularly relevant for the current discrimination grounds: digitalisation, poverty and standard of living, and mobility and transportation.

2.1.3. Addition of new features to the data hub

The content of the mapping tool translates into the **data hub**: the inventory was uploaded to the website and the sources that were already in the data hub since the IEDCB I project were updated. The different categories of information registered through the mapping tool (see Table 1) are displayed on the **information page of each source**, giving an idea of the content of the source.

Some **new features** were also added: searching for sources for the grounds that were the focus of this project (disability, health status, and physical characteristics), searching for sources based on keywords, and searching for sources containing intersectional data or analyses.

We noted up to 10 **keywords** per source, with the aim of being able to provide more details about a registered domain, about a registered ground, or about any general issues/phenomena addressed in a source. This allows the user to search for sources on more specific topics (via the 'Enter search terms' functionality or via the list of frequently used keywords) and to get a better idea of the topics covered in a source (e.g., a specific type of disability).

We also added a function for **intersectionality** to the data hub. Intersectionality is a perspective which states that the social situation of individuals is determined by the intersection of multiple personal characteristics[37] (see section 3.1 for more information on the application of this perspective in equality data). The new feature allows users to search for sources that analyse data in an intersectional way and to get an idea of which intersections are covered in a specific source. For example, if a source contained an intersectional analysis between the characteristics of skin colour and disability, we noted this.

2.2. Information gathering through crowdsourcing

We consulted **several key actors** throughout our work, both in order to conduct our project in a participatory manner and in order to call upon their expertise. This crowdsourcing was organised in different ways: through meetings with an advisory group, through the distribution of a survey, and through in-depth interviews. We tried as much as possible to capture different perspectives.

2.2.1. Meetings with the advisory group

At the start of the project, we set up an **advisory group** to guide us during the research. This advisory group was composed of people working for various government bodies (including the statistical institutions), civil society representatives doing work related to the current grounds, and academic staff. In total, this group consisted of about 35 (Dutch-speaking or French-speaking) members.

Three meetings with the members of this advisory group took place: in February 2023, in September 2023, and in January 2024. During these meetings, we presented the project and its progress and reflected together on certain topics.

As the process was based on the principle of participation, members were regularly invited to provide feedback and were also given the opportunity to provide their comments on an initial version of the written report.

2.2.2. Survey

An **online survey** was widely distributed in April 2023. The target audience consisted of employees of government bodies, members of civil society organisations, and academic staff, who were working on

[37] Rosa vzw. Intersectionaliteit.
Equality Challenge Unit (2017). Intersectional approaches to equality research and data.

the topic of equality data and/or on the topics of disability, health status, or physical characteristics. The survey was sent via mail to about 370 organisations/individuals, who were encouraged to further distribute the survey. In total, we received 50 responses, of which 30 were in Dutch and 20 in French. As the survey topic was quite specific, we aimed to reach approximately this number of respondents.

The responses were varied and came from respondents with quite **diverse profiles:** 24 of them were employees of government bodies of different types and different regions, 18 were members of civil society organisations advocating for different concerned groups, 7 were employees of universities or colleges, and 1 person worked for both a university/college and a civil society organisation.

Through open-ended questions, **several topics** were addressed in the survey (the full version in French or Dutch is available on this webpage):

- The ways in which respondents **use or collect** equality data
- The opportunity for respondents to share **relevant data sources** that we could add to our inventory, which allowed us to already identify about a hundred potential sources
- The **gaps and issues** identified by respondents in terms of equality data for disability, health status, and physical characteristics
- How **intersectionality** is addressed (or not) in the respondents' work
- The use and application of **self-identification** as a basic principle of the human rights-based approach to data
- **Other themes or comments** respondents wanted to bring to the table

We then analysed the answers to each question and summarised them in a synthesis. The results have been incorporated into this final report. The identified themes were used to shape the structure of this report (especially Chapters 3 and 4) and to guide our analysis of the in-depth interviews.

2.2.3. In-depth interviews

We consulted a total of **eight experts** through **in-depth interviews**: four individual interviews and two double interviews, three of which were conducted in Dutch, two in French and one in English. The experts in question were academic researchers as well as practitioners specialising in the fields of disability, chronic diseases, and mental health.

The interview guide included questions on the conceptualisation and definition of the grounds of disability, health status, and physical characteristics. We also included questions about gaps and issues in terms of equality data, about self-identification, and about intersectionality. Finally, we discussed the translation of these data into policy, as well as sensitivities and ethical issues related to these data. The semi-structured interview guide was flexibly adapted according to the expertise of each person we interviewed. When other themes emerged during the interview, these were also discussed in detail.

The interviews were recorded and transcribed and then analysed using the programme ATLAS.ti.

3. Gaps in Belgian equality data for disability, health status, and physical characteristics

In this section, we describe the gaps and issues that we identified in the Belgian equality data for disability, health status, and physical characteristics. We first analyse our **inventory** to get a quantitative overview of the data sources we selected and mapped: for which topics did we find a lot of data, for which topics did we find little or no data? We then turn to the results of our **crowdsourcing exercise** and discuss the gaps, issues, and concerns raised by the stakeholders we consulted.

3.1. Analysis of the inventory

At the time this report was written, the inventory of equality data included a total of 527 data sources. **232 of these sources cover at least one of the current discrimination grounds: disability, health status, and physical characteristics**. In this section, we take a look at which domains these sources cover, which intersectional data or analyses they include, what types of data they contain, in which languages they are available, and which geographical regions of Belgium they cover.

We limit ourselves in this section to the sources and gaps related to the current grounds: disability, health status, and physical characteristics. For the criteria already considered in IEDCB I (the so-called 'racial' grounds, religious and philosophical beliefs, sexual orientation, gender identity, gender expression, and sex characteristics), we refer to the table in the annex.

Table 2 shows for each discrimination ground the number of identified sources that contain data on the criterion in question, as well as how many of those sources contain data on specific domains.

It becomes immediately clear that we found quite a lot of sources of equality data for the ground of **disability** (187 sources in total). The domain that is by far covered most often in these sources is work, followed by health and well-being. There are also a fairly large number of sources on education and housing. In addition, the category 'Other domain' comes up quite frequently: we particularly found a lot of data on leisure and participation (e.g., sports, cultural activities, membership of associations or clubs), on social protection (e.g., benefits, financial support), on safety and experiences of (sexual) violence, and on access to (public) services. In contrast, it is notable that relatively few sources contain data on inequality based on disability in the domains of **justice and police**, **media**, **and digitalisation**.

The inventory also contains a fairly large number of sources related to the ground of **health status** (103). It should be noted, however, that 56% of these sources are also linked to the ground of disability. Indeed, in many cases the distinction was not very clear (e.g., in sources related to chronic diseases), leading us to register a source as pertaining to both grounds so that users of the data hub would not miss out on relevant sources. Still, we do see a slightly different pattern here compared to what we saw for the ground of disability. By far the majority of sources containing data on health status are related to the domain of health and well-being, and to a lesser extent to work, housing, and other domains (largely the same as for disability). Again, we found little data on justice and police, media, and digitalisation. There are also few equality data on **education** and **poverty** as a function of health status.

Finally, we found **very few sources on physical characteristics** (only 14 sources) and this was true for all domains. Most of these sources contained either data on the (mental) health of individuals and its possible relation to weight (hence the health and well-being domain recurring a few times), or data on complaints about discrimination. It is important to note that the low number of sources in the inventory is in part due to the selection criteria we applied (see section 2.1.1). For example, there exist some purely academic sources of data on discrimination based on physical characteristics[38]. Also, we often did not include international sources because they did not contain any specific data on Belgium.

TABLE 2. NUMBER OF SOURCES CONTAINING DATA ON A SPECIFIC DOMAIN (BY DISCRIMINATION GROUND).

Domains	Discrimination grounds		
	Disability	Health status	Physical characteristics
Work	74 (39.6%)	29 (28.2%)	4 (28.6%)
Education	43 (23.0%)	16 (15.5%)	5 (35.7%)
Health and well-being	64 (34.2%)	54 (52.4%)	8 (57.1%)
Poverty and standard of living	35 (18.7%)	19 (18.4%)	1 (7.1%)
Housing and living situation	56 (29.9%)	24 (23.3%)	3 (21.4%)
Mobility and transportation	21 (11.2%)	8 (7.8%)	2 (14.3%)
Justice and police	14 (7.5%)	8 (7.8%)	4 (28.6%)
Media	12 (6.4%)	4 (3.9%)	2 (14.3%)
Digitalisation	12 (6.4%)	2 (1.9%)	0 (0.0%)
Other domain	54 (28.9%)	24 (23.3%)	3 (21.4%)
No domain	22 (11.8%)	15 (14.6%)	1 (7.1%)
Total number of sources	**187**	**103**	**14**

(1) The percentages between brackets should be read column by column. They reflect the number of sources containing data on a domain, relative to the number of sources containing data on the discrimination ground. For example, 23% of the sources related to the ground of disability contain data on the domain of education.
(2) Because the same source can contain data on multiple domains, the sum per column is greater than the total number of sources for that ground (and greater than 100% for the percentages). Because the same source can also contain data on multiple grounds, the sum of the bottom row is greater than the total number of sources (232).
(3) For sources that cover multiple ground and domains, the fact that a ground and a domain are both covered in the source does not necessarily mean that the source actually looks at inequality based on that specific ground in that specific domain (although this is usually the case). The numbers in this table thus represent an upper bound for the actual number of sources containing data on the *intersection* of a ground and a domain.

We next consider which **intersectional data or analyses** regularly appear in the identified sources (see the box 'Intersectionality in equality data'). Here, we limit ourselves to the sources on disability and health status; the few sources that were identified for physical characteristics contain almost no intersectional analyses. **Age** is regularly (in 20-25% of the sources) crossed with the grounds of disability

[38] See, for example, Baert, S. (2018). Facebook profile picture appearance affects recruiters' first hiring decisions. *New Media & Society, 20*(3), p. 1220-1239.

or health status. The same goes for data crossed with **gender**[39]. In addition, some sources contain data on the intersection of disability with nationality or origin, education level, or family composition. However, intersections with other grounds occur in fewer than five cases or not at all in the inventory.

Moreover, this rarely involves a **'full-fledged' intersectional analysis**. The majority of these sources contain a mere crossing of variables (e.g., a table that is disaggregated by both disability and gender). While such data can allow one to conclude that one subgroup is more disadvantaged than another, they do not provide insight into the precise nature of that situation and its structural causes. Only a few of the identified sources contain an in-depth, qualitative analysis of the situation of persons at a specific intersection.

INTERSECTIONALITY IN EQUALITY DATA

The term 'intersectionality' was introduced by Kimberlé Crenshaw in 1989. This perspective argues that the social positions of individuals are determined by **multiple personal characteristics**[40]. Indeed, 'intersections' of several characteristics can create a **new situation** that is qualitatively different from the simple sum of the individual characteristics. For example, women with disabilities may experience unique forms of discrimination or inequality because they find themselves at a specific intersection between gender and disability. If we focus on only one characteristic at a time (e.g., only gender or only disability), these experiences remain invisible.

An intersectional perspective is also important for **equality data**. If equality data look at only one ground, we may not get a complete picture of the discriminations that people experience. The final report of the project IEDCB I therefore devoted a chapter to intersectionality[41]. This chapter discussed how intersectionality can be taken into account in **various ways** in sources of equality data.

In some cases, it involves a **'full-fledged' intersectional analysis.** The situation of individuals who find themselves at a specific intersection is analysed in depth, focusing on the unique nature of the situation, structural power balances, and the causal links between the two. Qualitative research in particular lends itself well to this type of intersectional analysis.

Many quantitative sources contain a **more minimalist version** of intersectional data or analysis. Different grounds are crossed with each other, providing data for each subgroup (e.g., data on persons with disabilities with the tables disaggregated by age). Based on this, however, it is difficult to estimate whether this is has to do with qualitatively different situations. Regression analyses looking at interaction effects between several characteristics can partially address this, but even these provide little insight into the structural causes of inequality.

How diverse are the sources in terms of the **research methods** used? By far the most data in the inventory come from surveys (130 sources), administrative data (96 sources), and qualitative research (62 sources). To a lesser extent, we also found data coming from diversity monitoring (18 sources) and complaints data (15 sources). Interestingly, only 6 sources report the results of **discrimination testing**

[39] We have opted to generally use the term 'gender' when noting intersectional data or analyses, even if the original source used the term 'sex'. Indeed, many of these sources do not measure biological sex but rather how the person identifies. Only when the source explicitly focused on biological sex did we use the term 'sex'.
[40] Rosa vzw. Intersectionaliteit.
[41] Unia (2021). Final report 'Improving equality data collection in Belgium', Chapter 5: Intersectionality (available only in French or Dutch).

(with the important side note that purely academic sources using discrimination testing were not included in the inventory).

The **language of sources** is an important factor for making equality data accessible to as many users as possible. 147 of the sources are available in Dutch (of which 71 only in Dutch). About the same number of sources are available in French (144, of which 65 only in French). 75 sources are available in both languages. Finally, 36 sources are (also) published in English and 18 in German. Thus, although many sources are published in multiple languages, the majority are still available in only one language.

Finally, it is clearly relevant for policymakers to have data for **specific geographical areas**. Almost half of the data sources contain data for the whole of Belgium (102 sources). We also found a lot of data specifically on Flanders (88 sources). For Wallonia and Brussels, there are slightly fewer data (57 sources in both cases). Finally, the inventory contains very little data specific to the **French Community** (23 sources) or the **German-speaking Community** (13 sources).

3.2. Gaps and needs identified through crowdsourcing

We will now address the gaps and needs identified through the crowdsourcing exercise. Where relevant, we will link these results to the results of the inventory.

3.2.1. Issues with definitions and data comparability

The crowdsourcing revealed that **definitions** that are not clearly **comparable** between one another lead to a lack of clear and reliable data.

A first striking example cited during the crowdsourcing is the fact that numerous different definitions of '**disability**' are used in data sources. We discuss the issues surrounding definitions of disability and related concepts, as well as their operationalisation in the context of data collection, in more detail in section 4.1 ('Thematic focus: defining and measuring disability') of the 'Operationalisation' chapter.

The experts we encountered also cited **'burnout'** as another example of this problem: the many data that exist on the topic of burnout are difficult to compare because they are collected based on many different definitions.

Another example mentioned during the crowdsourcing is the need for a precise definition of ***actual* accessibility**, which - once established - would allow one to systematically measure the achievement of this principle. In other words, stakeholders expressed a desire to collect data on the possibility to truly participate in society, which includes data on access to work, to media, to hobbies, etc. By referring to the term *actual* accessibility, the consulted actors meant, for example, trying to measure whether a person with a motor disability, who receives assistance at home during specific hours, has the same access to the labour market as a person without a disability, when the working hours offered by the assistance service do not allow one to fulfil the obligations associated with work life.

Based on several examples from the crowdsourcing, we conclude that in order to collect data in an optimal way and to ensure their comparability, these data should preferably be collected according to **the same definition**. At a minimum, the definitions that are used should be **explicitly stated** so that any users of the data can assess their comparability.

3.2.2. Lack of complete, reliable data on disability

Regarding equality data on **disability**, the crowdsourcing revealed several gaps: some data exist but are incomplete or biased, and others are lacking. At the same time, it became clear that many future data collection efforts are planned. In this section, we address various elements that emerged during the crowdsourcing.

During the in-depth interviews, some experts told us that **many data already exist** without them being specifically collected, because they are **not explicitly seen as data**. This is the case, for example, for the waiting lists that exist for certain facilities, such as day or night centres, centres for special education, or centres for diagnosis and care. **Other data already exist but are not available**, such as data on certain types of requests or applications by municipality. According to our advisory group, making these data available would make it possible to study whether there is a link between the different situations.

Discussions with the advisory group and the experts emphasised that **barriers in the person's environment** are an essential element in the UN Convention's definition of disability (see section 1.3.1). The experts pointed out that in data collection, however, the variable 'disability' is mainly approached from a medical perspective. Not enough data are collected in which the variable 'disability' is understood as a disability arising from the interaction with the environment (that creates barriers to participation in society). A need for data on situational elements (barriers in a particular environment) was pointed out, for example high-quality data on the accessibility of buildings or websites for persons with disabilities.

Moreover, people indicated during the crowdsourcing that disability is often still conceptualized as a **uniform situation** in data collection efforts, without measuring **the nature and the degree of the disability**, which differ from person to person. The discussions showed that it is very important, however, to measure the various realities of persons with disabilities so that the specific situations of each subgroup can be taken into account. Methods that allow for this are discussed in Chapter 4.

The crowdsourcing also revealed that **certain specific groups** are insufficiently included in data collection. One example that was cited is that of persons living in facilities for collective living (e.g., institutions), as these groups are often not approached for surveys. Another example of a group that is not sufficiently covered in data collection is children with disabilities. In the domain of education, for example in schools, discrimination still remains largely under-reported. Finally, the advisory group mentioned the example of autistic persons or persons with severe intellectual disabilities, for whom there is also insufficient attention in data collection efforts. The crowdsourcing exercise suggested several ways this situation could be improved: firstly, more attention should be paid to specific groups that are regularly excluded from data collection, and secondly, research methods should be adapted in order to survey the persons in question through appropriate working methods.

The consulted stakeholders, and the experts in particular, stressed the need to **quantify the truly inclusive participation of individuals**, and to reflect on how to develop appropriate policies and tools to realise this true participation. Several initiatives are being launched to create an inclusive society, but according to the experts, there is still a long way to go. They stressed the need to have data that would allow to measure the progress being made. To achieve this, several avenues were suggested. For example, it is necessary to know the figures on funding for institutions and to be able to compare them with those for effective participation and inclusion. We also need figures on the number of people living in institutions, the waiting times to receive personal assistance budgets, and the waiting times to get a place in a facility. We also need surveys on institutionalisation experiences and on independent living.

The crowdsourcing additionally underlined the importance of collecting data on **reasonable accommodations**[42]. Actors called for making it possible to quantify how often these are requested, how often they are granted and in what contexts. These data would allow to evaluate the effectiveness and standardisation of reasonable accommodation procedures and to analyse which needs are currently not being met. In this regard, experts indicated that making existing data available would already be a step forward.

The survey respondents also mentioned the importance of collecting data on **non-access to and non-take-up of rights,** which would provide more precise information on the situations in which this phenomenon occurs and the profiles of the people concerned.

The **domain of work** was mentioned in the crowdsourcing as being incomplete in terms of data collection, although Table 2 showed that this domain is covered by many of the sources in our inventory related to these grounds. 39.6% of the sources for the ground of disability contain data on the work domain. However, according to the crowdsourcing, the current data do not allow for a complete picture of the situation of persons with disabilities on the labour market, neither in the private sector nor in the public sector. An example of this is that there are not enough data on persons working in companies for adapted work and on how many of them subsequently proceed onto the general labour market.

Finally, during the crowdsourcing people pointed out the need to include the ground of disability more broadly in data collection, i.e., to work on **including a disability variable in data collection by default**, naturally while still paying due attention to the goal of the data collection.

3.2.3. Lack of complete, reliable data on health status and physical characteristics

Our advisory group pointed out that the scarcity of data on **health status** can be explained, among other things, by the fact that medical records are of course confidential, which makes data collection more complex. Discussions with the group indicated that this gap is also explained by the lack of a **harmonised definition** of what is understood by health status. This leads to the available data not being comparable, as the methodology for the definitions used is not explicitly explained, and even if it is, the methods are not comparable. As such, very few data are collected that focus specifically on the criterion of health status. Our inventory confirmed that it is not easy to make this distinction, as data on health status often also seem to fall under the definition of disability.

Regarding the ground of **physical characteristics**, it is very clear, both from the crowdsourcing and from our inventory, that insufficient data are collected for this ground. It was mentioned in the crowdsourcing that there is no standardised measurement tool for collecting these data and that such a tool should be developed.

3.2.4. Lack of intersectional data

Our discussions with the advisory group and the various experts clarified **which intersectional data are lacking** for the current grounds.

Two main intersections were highlighted during the crowdsourcing: the intersection with gender and the intersection with age. This is consistent with the fact that the UN Convention on the Rights of

---

[42] Reasonable accommodations are appropriate measures taken to address a specific need in a concrete situation, to enable a person with that need to access and participate in a particular environment, unless these measures impose a disproportionate burden on the person who has to provide them. See Unia. [What are reasonable accommodations?]

Persons with Disabilities calls for special attention to the situations of women and children with disabilities[43].

Regarding the intersection with **gender**, stakeholders mentioned a lack of data on access to the labour market for women with disabilities. In fact, a special module (linked to the Labour Force Survey) was conducted in 2011 under the coordination of Eurostat, which focused on the labour force participation of persons with disabilities. However, this exercise has not been repeated since then. Another gap that was cited concerns data on the specific experiences of women with a disability and a migration background. Although we saw in our inventory that the intersection of gender and disability is more common in data than some other intersections, the crowdsourcing revealed that even for this intersection there is still insufficient data available.

For the intersection with **age**, and in particular concerning the experiences of children with disabilities, there is too little data on children with learning disabilities and on children with neurodevelopmental disorders, according to the crowdsourcing.

The crowdsourcing also revealed a great lack of intersectional data on **intersections between the current grounds and other grounds**, including social origin or situation (income, poverty, level of education), national or ethnic origin, skin colour, language, family situation (marital status, whether one is single, how many children one has), and sexual orientation. However, several experts have stressed the importance of these intersectional data to understand the situations of the people involved. For example, having a disability can have very different consequences depending on the income of the person or their family, and these consequences are often more severe for people in more vulnerable situations.

In some cases, however, it is possible to obtain micro-data based on a relatively simple request, e.g., to Statbel, with which certain intersections can be analysed, when this is not possible based on the published survey data. For example, based on the Labour Force Survey, it is possible to cross the GALI variable (which contains information on disability[44]) with gender and nationality of origin. This possibility of working with micro-data based on existing survey data is one way of filling certain gaps in terms of intersectional data.

We can clearly conclude from the crowdsourcing that more intersectional data should be collected, for a wide range of grounds. Without these data, complex situations cannot be understood in their entirety.

3.2.5. Limited availability and accessibility of data

The demand for **better accessibility of data** was clearly reflected in many exchanges during the crowdsourcing. Data are hard to find and are not always freely accessible to the general public.

The crowdsourcing highlighted that - although many data still need to be collected - a lot of data already exist, but are often **inaccessible** and **not comparable**. They are available to local authorities for example, such as data on the number of people receiving some form of assistance. However, these data are not published and are therefore not accessible. The advisory group indicated that work should be done to make these data available, even if this would entail anonymising the data and updating them regularly.

[43] United Nations (2006). Convention on the Rights of Persons with Disabilities, art. 6 and 7.

[44] GALI stands for *Global Activity Limitation Indicator.* More information on this can be found in section 4.1.2.

Another accessibility barrier that was cited by the advisory group is that of **language**. Sources are often not available in certain languages, such as German or English. Admittedly, translating existing data into these languages would require a workload that is often unfeasible for the producer of the data.

Data accessibility for persons with disabilities is also an issue that emerged in the crowdsourcing. This lack of accessibility can be due to the use of difficult language (especially for persons with intellectual disabilities), but also to other issues, such as in the case of websites that do not meet digital accessibility standards (causing problems for screen reading software, etc.) or in the case of visual or auditory content that is not accessible to persons with sensory disabilities. The question of **data understandability** also arises. Data are often presented in their raw form, when they could be better framed and interpreted in order to make it easier for readers to extract specific information.

The crowdsourcing also addressed the fact that some databases contain information on disability or on health status, but that it is not possible to **link these to other variables** in other databases. This makes it impossible to link these variables even though they cover the same population.

Several practitioners have informed us that being aware of the relevance of the data they hold is not always obvious. To address this, there is a need for a **framework and guidelines** on how to handle such data correctly. This is the case, for example, for all data on requests that are submitted to facilities that provide services to persons with disabilities (waiting lists, requests for transport and housing, etc.).

3.2.6. Distribution of jurisdiction and lack of translation of data into policy

Equality data are useful to objectify facts, but the crowdsourcing highlighted that they are often **not followed by political decisions**. The experts explained that Belgium's institutional landscape makes it difficult to implement cross-cutting and global policies, as competences are divided between different levels of government. The advisory group indicated that differences between the realities of each region are often overlooked due to the lack of figures to map these differences. The distribution of competences among different entities also makes it difficult to approach certain issues in an intersectional manner, even though they are linked to the intersection between different grounds and thus need to be addressed in a cross-cutting manner.

The **comparability** of data (even those that are available) is also **complex** because the criteria and the variables that are used are not the same, for example in terms of when the government considers someone to have a disability. One example is that of the labour market: several sources contain data on persons with disabilities, but there are almost no sources with central data for the whole of Belgium (with the exception of certain sources from Statbel).

During discussions with our advisory group and in the survey, reference was made to the need to have **more local data** (e.g., by municipality). While particular attention should be paid to anonymising such data and to checking that the number of individuals per municipality is not too low in order to ensure confidentiality, these data are essential as many policy measures are determined at the local level. Some data at the local level already exist, for instance in the Municipality and City Monitor[45] and in the database of the InterMutualist Agency (IMA)[46], but they remain rather limited in terms of variables linked to the current grounds (disability, health status, and physical characteristics).

Finally, the advisory group reported that the **German-speaking Community** find itself in a very specific situation. Residents often turn to Germany and the Grand Duchy of Luxembourg for help and access to certain services that are not available in German within Belgium. This is the case for health-related

[45] Website of the Municipality and City Monitor.
[46] L'Agence InterMutualiste. Données de santé.

services that require a dialogue of high quality. For example, it is important to realise that there are no interpreter services for German Sign Language available in the German-speaking Community. In sum, if the aim is to get an accurate picture of the use of certain services or of certain needs, this will lead to an underestimate because the requests made within Belgium do not correspond to the reality of the needs.

4. Operationalisation of discrimination grounds in equality data

To arrive at data that can say something about inequality or discrimination based on disability, health status, or physical characteristics, it is crucial to be able to **measure** these concepts. For example, the data should include a variable that makes a distinction between **persons with disabilities and persons without disabilities,** allowing us to compare the situation of these two groups.

A key question, therefore, is how to operationalise these grounds of discrimination. **'Operationalisation' refers to the process of translating abstract concepts (such as disability) into concrete, measurable variables**. Of course, this is strongly related to how we define the concept in question. But even if we use the same definition, a concept can be operationalised in different ways. In order to collect reliable equality data and interpret them correctly, we should always consider the operationalisation of the relevant discrimination grounds.

For defining (and operationalising) disability and health status, it is clear that tension may exist between how individuals perceive their **individual reality**, what **diagnosis or diagnoses** they may receive, and whether their disability is **officially recognised** by government bodies (e.g., the FPS Social Security). There is rarely a perfect overlap between these different aspects. In addition, the concept of disability is also strongly linked to the **broader social context**: the barriers that a person experiences depend on how well or poorly adapted society is. In the definition of physical characteristics, **social aspects** likewise play a role: whether one is at risk of discrimination based on a particular characteristic depends on whether there is a societal stigma connected to it.

In this chapter, we look at how the current discrimination grounds can be operationalised in (equality) data. In the first part, we give a comprehensive overview of the different ways in which the ground of disability is measured (and to some extent the ground of health status, as it is often measured together with disability). In the second part, we consider a complex issue that seems to touch on each of the current grounds and about which we saw much confusion throughout the project: the relationship between the concept of weight discrimination on the one hand, and the different perspectives on the term obesity on the other hand.

4.1. Thematic focus: defining and measuring disability

In the past, people often used very strict, medical definitions of disability. Nowadays, the **definition from the UN Convention** is increasingly seen as the reference, which is a good evolution. This definition is now accepted in national and international case law and unambiguously identifies the target group of human rights and anti-discrimination legislation related to disability. At the same time, the operationalisation of this definition brings with it clear challenges in the context of data collection. The definition is very broad and covers many different situations and individuals, all with different experiences and needs. Consequently, during the crowdsourcing, concerns were raised that certain groups (e.g., persons with severe intellectual disabilities) may become even more invisible if we consider 'persons with disabilities' as one group when collecting data. Indeed, knowing whether someone has a disability is not enough: data should also allow for **differentiation within the population of persons with disabilities** (e.g., by including a variable that reflects the nature of the disability).

In this section, we discuss in detail the ways in which disability and related concepts can be operationalised in equality data. We start by giving an overview of the **main concepts**: what exactly do we want to measure? We then cover **four ways in which data are often collected**: (a) through questions asked in surveys, (b) through the use of information collected in the context of administrative procedures, (c) through the use of other proxy variables, and (d) through self-identification. For each method, we discuss which concepts can be measured with that method and what the **advantages and disadvantages** are.

For this part, we rely on information provided during the crowdsourcing exercise and on a review of relevant literature. In addition, the report of the Working Group on Statistics (coordinated by the FPS Social Security) provided a very valuable source of information[47].

4.1.1. Overview of key concepts

For each of the methods we discuss in the following sections, it is important to ask the question: What exactly is being measured? For the current purposes, we focus on the concepts below:

- The **presence of a disability**. Does a person have a disability or not? A variable that measures this concept is needed to compare the situation of persons with and without disabilities. In most cases, this variable will naturally be a proxy[48], as few data sources operationalise the presence of disability in full accordance with the definition in the UN Convention.
- The **nature** and the **severity** of the disability. Measuring these concepts allows for differentiation within the population of persons with disabilities. If the data contain such variables, we can better map the situation of specific groups.
- The **degree of autonomy**, **the degree of required care**, and **specific needs in terms of support**. These concepts are mainly measured in the context of administrative procedures for determining whether to assign support (financial or otherwise) to a person. However, they are also very relevant in the context of equality data because they say something about the barriers to participation for people with disabilities and the changes required to remove these barriers.
- **Health** or **health status**. These concepts are also regularly measured in data collection related to disability. They refer to a person's health in general or to the presence of health problems, regardless of whether this causes participation problems for the person.
- **Specific conditions**. Finally, sometimes the presence of specific conditions or diseases is also measured, regardless of whether they cause participation problems. Measuring this concept allows for differentiation within the broad concept of health status.

4.1.2. Survey questions

Several large-scale surveys include questions that measure one or more of the above concepts. Operationalising these concepts through surveys questions in order to create equality data has some **general advantages**:

- Surveys assess the presence of a disability relatively directly and do not have to rely on an external indicator (e.g., whether or not someone is officially recognized as a person with a disability). This allows one to use a **broad definition** of disability and to identify individuals not known to official bodies.
- Moreover, the same questions are often asked across different surveys (and their international versions), which improves **comparability**.
- Finally, the individuals themselves provide the data on their personal characteristics, which is in line with one of the human rights principles we cited in the introduction: **self-identification**. In section 4.1.5, we reflect on what full-fledged self-identification means.

[47] SPF Sécurité sociale (2023). Groupe de travail statistiques - CIM Handicap - Note finale.
[48] A proxy is a variable that substitutes the variable one is actually interested in, when that variable is impossible or difficult to measure. The proxy approximates the variable in an imperfect manner. For example, a proxy of 'disability' can be whether someone is recognised by the government as a person with a disability.

However, there are also **disadvantages** to operationalising the concepts through survey questions:

- The **context** in which questions on disability are asked must be taken into account. If the entire survey focuses on the work context, respondents will mainly think of participation problems at work when answering a question on disability, whereas they may answer differently in a survey on another topic. This can partly be overcome by asking questions about disability near the beginning, making the answers more comparable across surveys. At the same time, it will often be desirable for the individual survey if respondents keep the context in mind when answering, in order to gain insight into the barriers people experience in that specific context.
- In addition, people have to **self-report** whether or not they have a disability or condition. This carries the risk of **over- or under-reporting**. In the crowdsourcing, for example, we encountered resistance to the idea of letting people report whether they have autism without linking this to an official diagnosis, as this could lead to overreporting. Conversely, it was mentioned that persons considered to have mild intellectual disabilities do not always see themselves as having a disability. An additional risk of underreporting arises from the fact that existing questions are not always well suited to measuring mental health problems, as we will see below.

We now discuss a number of specific questions that are often included in large-scale surveys. By far the best known instruments in the context of disability are the *Minimum European Health Module* (which includes the *Global Activity Limitation Indicator*) and the *Washington Group Short Set on Functioning*.

The ***Minimum European Health Module*** **(MEHM)** consists of three questions[49]:

1. How is your health in general?
   Very good / Good / Fair / Bad / Very bad
2. Do you have any longstanding illness or health problem?
   Yes / No
3. For at least the past 6 months, to what extent have you been limited because of a health problem in activities people usually do?
   Severely limited / Limited but not severely / Not limited at all

Questions 1 and 2 measure **health status** independent of any participation problems. Question 3 is known as the ***Global Activity Limitation Indicator*** **(GALI)** and is sometimes asked by itself. This GALI question combines several elements: the presence of a health problem, its impact on participation, and the long-term nature of this impact. As these elements correspond quite well to those in the UN definition of disability, the GALI is often used as a proxy for the **presence of a disability**[50]. Moreover, because it offers a choice between 'Severely limited' and 'Limited but not severely', this question also measures the **severity of the disability**.

There are clear **advantages** to using the GALI question:

- The presence of disability and its severity are measured in an **unambiguous, concise manner**.
- The question is used - partly due to coordination at the European level - in a great number of surveys, including the Statistics on Income and Living Conditions, the Labour Force Survey, the Health Interview Survey, the Household Budget Survey, the Survey of Health, Ageing and Retirement in Europe, the European Quality of Life Survey, etc. This is a great asset for **comparability** between surveys, between countries, and over time.

[49] Eurostat. Glossary:Minimum European Health Module (MEHM).
[50] Eurostat (2020). Standardised key social variables: Implementing guidelines, p. 66-68.

- The question seems to show good **validity and reliability**[51].

However, we also note some **concerns** regarding the use of the GALI:

- Although the question corresponds fairly well to the UN definition, it does not explicitly refer to **barriers** in the environment. Rather, this aspect is implicitly contained within the reference to a limitation in 'activities people usually do', which leaves room for interpretation.
- It is also possible that people interpret 'health problem' in a more strict manner than the 'impairments' that the UN definition refers to, which may lead to underreporting. In line with this, research suggests that people's responses to the GALI are less strongly linked to their **mental health problems** than to their physical health problems[52].
- The question is rather complex. It is therefore sometimes divided into **two questions**, after which the answers to both questions are combined[53]. However, this choice differs between surveys, which can decrease the comparability.
- The question is not systematically **translated** in the same way to other languages. The Dutch and French translations that were included in the original versions of this report were the ones used in Sciensano's Health Interview Survey. In other translations to Dutch or French, parts of the question are sometimes worded differently[54]. The answer options are also often translated (or back-translated) differently[55]. The differences may seem subtle, but they can lead to problems of comparability. An official translation would be a good thing.
- The question has been developed for persons aged 15 years and above, so it cannot simply be used for data collection in **children**.

The **Statistics Flanders survey**[56] (a large-scale survey specific to the Flemish region) asks questions similar to the GALI question:

1. Do you have any long-term disease(s), condition(s) or disabilit(y)(ies)?
   Yes / No
2. How often are you restricted in your daily activities by these disease(s), condition(s) or disabilit(y)(ies)?
   Rarely or never / Occasionally / Constantly

However, these questions differ fundamentally from the standard GALI question. For example, the first question does not provide any differentiation between health status and disability. Moreover, no reference point of 6 months is mentioned, leaving people to estimate themselves what 'long-term'

[51] Validity refers to the extent to which the question accurately measures the underlying concept (in this case, disability). Reliability refers to the extent to which the question consistently leads to the same results.
Van Oyen et al (2018). Measuring disability: a systematic review of the validity and reliability of the Global Activity Limitations Indicator (GALI). *Archives of Public Health, 76*(25).
[52] Van Oyen et al (2018). Measuring disability: a systematic review of the validity and reliability of the Global Activity Limitations Indicator (GALI). *Archives of Public Health, 76*(25).
[53] Eurostat (2020). Standardised key social variables: Implementing guidelines, p. 66-68.
[54] 'Activities people usually do', for example, is replaced by phrases such as 'your daily activities' (Labour Force Survey) or 'your activities' (the MEqIn survey). Moreover, the Belgian version of the Labour Force Survey does not refer to a 'health problem' but to 'a disability, a disorder or a disease' (to avoid only thinking of diseases when reading the word 'health problem'). In this case, the word 'disability' is explicitly mentioned.
[55] The answer option 'severely limited' is back-translated in English as 'yes, strongly', 'yes, strongly limited', 'yes, severely', and 'yes, severely limited', among others. The response option 'limited but not severely' is back-translated as 'yes, to some extent', 'yes, limited', and 'yes, a bit limited', among others.
[56] Statistics Flanders. SF survey. Note that the version here is a translation from Dutch rather than the English version of the question, as the full English survey is not publicly available. There may be small differences.

means. This may have an impact, all the more so because the second question (which assesses participation problems and their severity) is only asked if people answer 'Yes' to the first question. Finally, severity is assessed on a different scale (how often). Because of these differences, it is difficult to know whether data from the Statistics Flanders survey are comparable to other Belgian data based on the GALI.

GALI: TAKE-AWAY

Even though the GALI question does not measure all elements of the UN definition, this widely used and validated instrument can certainly be considered when collecting equality data that require a variable measuring the presence of disability (and to some extent its severity).

However, this requires careful reflection with regard to the exact translation and implementation, or preferably working with an existing version.

A second important tool is the ***Washington Group Short Set on Functioning* (WG-SS)**. It contains **six questions that assess difficulties with seeing, hearing, mobility, cognition, self-care, and communication**[57]. The underlying idea is that difficulties with these basic actions will often lead to participation problems in an unaccommodating environment. As such, the questions themselves do not refer to participation problems or environmental barriers. The questions take the following form[58]:

1. Do you have difficulty seeing, even if wearing glasses? Would you say...
   No difficulty / Some difficulty / A lot of difficulty / Cannot do at all

The WG-SS is used less systematically in surveys than the GALI. However, the triennial 'Health' module of the Statistics on Income and Living Conditions does include the questions. It has also been suggested as a measurement tool in the context of the Sustainable Development Goals (SDG) related to discrimination, with the aim of being able to disaggregate data on discrimination by disability[59]. In the context of the Belgian evaluation of the UN Convention, both the UN Committee and Unia also recommended using this tool in data collection[60].

Indeed, the WG-SS has several key **advantages**:

- The WG-SS measures not only the presence of a disability[61] and its severity, but also the **nature of the disability**.
- The developers of the tool centralise **clear and detailed documentation** on their website[62]. This helps to use the tool correctly and make sure that the results are as comparable as possible. In addition, the website offers several official translations, and even unofficial translations are kept track of by the developers.

---

[57] Washington Group on Disability Statistics (2022). WG Short Set on Functioning (WG-SS).
[58] There is also an official translation available in French, see Washington Group on Disability Statistics. Translations of WG Question Sets. There is not yet an official Dutch translation.
[59] Office of the High Commissioner for Human Rights. Guidance Note for Implementation of Survey Module on SDG Indicator 16.b.1 & 10.3.1 (Discrimination).
[60] Committee on the Rights of Persons with Disabilities (2019). List of issues prior to the submission of the combined second and third periodic reports of Belgium.
Unia (2024). Parallel report to the Committee on the rights of persons with disabilities, p. 35.
[61] The manual for the WG-SS proposes four possible thresholds for determining whether someone has a disability. See Washington Group on Disability Statistics. WG Short Set on Functioning (WG-SS) Syntax.
[62] The Washington Group on Disability Statistics

- A specific version has been developed for **children** aged 2 to 17 years: the *WG/UNICEF Child Functioning Module*[63].

There are also some **reservations** to be taken into account. The WG-SS of course only surveys six impairments in functioning and hence does not allow very high **specificity** in measuring the nature of disability. Moreover, some persons with disabilities may **not be identified** with the WG-SS. Indeed, a person may experience participation problems stemming from a different impairment than the six being surveyed. This is the case, for instance, for persons with mental health problems, such as anxiety and depression. These are however covered by the extended form of the tool (the *WG-SS Enhanced*), which includes additional questions about problems with affect (anxiety and depression) and upper body motor skills[64].

WG-SS: TAKE-AWAY

The key advantage of the WG-SS is that it measures the nature of disability, allowing the situation of specific groups to be analysed when the instrument is used in the collection of equality data. At the same time, only six specific impairments are measured.

Thus, if we want to use a broad definition of disability, it is advisable to use the WG-SS Enhanced and/or combine the WG-SS with a general question such as the GALI.

Besides the GALI and the WG-SS, more specific questions are regularly used.

For example, the Health Interview Survey[65] asks about (a) **specific diseases and conditions** in the past 12 months, (b) **long-term limitations** in various functions (similar to the WG-SS, but in more detail), (c) **difficulty with daily activities** (only for people over 50 years old), and (d) the **help** one would need in order to do them. As such, almost all of the concepts we distinguished at the beginning of this section are measured. In new efforts to create equality data that include these variables, an interesting avenue could be to make maximum use of existing data from the Health Interview Survey, or to take inspiration from questions included in the Health Interview Survey when collecting new data.

The Labour Force Survey also contains an additional question, which gauges **specific support needs in the context of work**[66]. Such a question could certainly be useful in other domains to gauge how the environment could be better adapted.

Finally, in the context of equality data, it is interesting to look at questions that do not measure disability as such, but rather **discrimination based on disability**. For example, the EU-MIDIS II survey asks whether people have ever felt discriminated against for various reasons, including disability[67]. Depending on whether one wants to assess general inequality or more specifically discrimination, it may be relevant to include such a question.

---

[63] Washington Group on Disability Statistics (2020). WG/UNICEF Child Functioning Module (CFM).
[64] Washington Group on Disability Statistics (2022). WG Short Set on Functioning - Enhanced.
[65] Sciensano. Health Interview Survey.
[66] SPF Sécurité sociale (2023). Groupe de travail statistiques - CIM Handicap - Note finale.
[67] European Union Agency for Fundamental Rights (2017). Second European Union Minorities and Discrimination Survey - Technical report.

USE OF SURVEY QUESTIONS: TAKE-AWAY

Overall, we see that many **survey questions** already exist that operationalise the presence of disability and related concepts. For new initiatives regarding the collection of equality data, the above overview can provide guidance for selecting the most appropriate questions, taking into account important considerations.

As for the existing surveys, progress can still be made in terms of (a) more systematically measuring the **nature** of the disability, (b) **coordinating** which questions are asked in Belgian surveys, and (c) further **standardising** their precise translation and implementation[68].

4.1.3. Collection of information for administrative purposes

Many data are also collected for **administrative purposes (mainly related to social protection),** which contain information on disability and other related concepts.

For example, the Directorate-General for Persons with Disabilities of the FPS Social Security has data on the persons who are officially recognised by them as persons with disabilities and on the persons who are entitled to the various allowances and benefits they grant (such as the integration allowance or the parking card). They also have all of the information that they collect in order to decide on the recognition and granting of those rights, including data on the degree of autonomy and the specific pathology (nature of disability)[69].

This example illustrates that we can distinguish two forms of administrative data:

- **Data indicating whether a person is recognised as having a disability or has been granted a particular right.** This type of data can (and is) used as a proxy for the **presence of a disability**, as it make a clear distinction between people who have been recognized/granted a right versus those who have not. It should be stressed here that it inherently concerns a **proxy**: there is never a one-to-one relationship between the elements in the UN definition and the criteria that are used by an authority for disability recognition.
- **Data collected by the authority in order to decide whether to recognise or grant a right**. These data are much more detailed and may include information on the **severity** of the disability, the **nature** of the disability, the **degree of autonomy** and/or **specific support needs** (depending on the authority and the right in question). Here, it is more difficult to gain insight into exactly what data are kept and in what form. The report of the Working Group on Statistics provides an overview of various government agencies[70], but it remains unclear how systematically these data are kept and in what form they could be made available for analysis.

By themselves, these administrative data can be used to get **the best possible picture** of the population of persons with disabilities in Belgium and their characteristics, as well as how often a particular right is granted, how many people rely on a particular service, etc. This constitutes important contextual information for equality data.

[68] For the national surveys, we realise that the room for manoeuvre is somewhat limited as these are partially coordinated at the international level.
[69] SPF Sécurité sociale. Statistiques Personnes handicapées.
SPF Sécurité sociale (2023). Groupe de travail statistiques - CIM Handicap - Note finale.
[70] SPF Sécurité sociale (2023). Groupe de travail statistiques - CIM Handicap - Note finale.

However, to effectively analyse **inequality** (e.g., in the work domain), these data need to be **linked** to other relevant data (e.g., data indicating whether one is employed or not). This can be done, for instance, via the Crossroads Bank for Social Security[71]. The data of some agencies are already integrated in this central database, the data of other agencies are not (yet) integrated. The Working Group on Statistics has plans to further increase the centralisation of data via the Crossroads Bank.

For an **exhaustive overview of which agencies collect which data,** we refer to the report of the Working Group on Statistics, which compiles this information for various federal and regional agencies[72]. Below, we do discuss the general advantages and disadvantages of using these data.

Using existing administrative data to create equality data has significant **advantages**:

- The data can be linked to a whole range of **other data** to assess inequality, without requiring additional data collection.
- The data also contain information about concepts such as the **severity and nature of the disability**, often in much more detail than survey data.
- Finally, the data do not depend on the person's own estimation and are therefore seen as **more objective**, although this can also be called into question (see below).

Unfortunately, there are also many **problems** with the available administrative data in Belgium:

- Competences, and therefore data, are spread across **many different agencies**.
    - Each agency starts from a different logic and objective when determining their **criteria for recognition**, which in addition may also vary from one right to another. As a result, the 'presence of a disability' may be based on completely different criteria in one case versus another. Sometimes criteria unrelated to disability, such as income or family composition, also come into play.
    - Depending on **age**, people are eligible for different rights, with children and older persons generally being covered by different systems[73].
    - Some competences have been transferred from the federal level to the regions in recent years, interrupting data **continuity**.
    - This fragmentation affects the **comparability** of the data. Moreover, it is highly challenging or even impossible to identify the entire population of persons with disabilities in Belgium using the same variable.
- Moreover, **non-access to and non-take-up of rights** is known to be a major problem[74]. If people do not apply (because they do not know that they are entitled to a right, because they do not succeed in submitting the application, or because they think they will not meet the conditions), they naturally do not appear in the data.
- There is often a **lack of documentation and transparency** on what data are kept and how this is done. For example, it is not always clear how systematically data are kept on rejected

[71] Crossroads Bank for Social Security

[72] FPS Social Security (DG Persons with Disabilities), Flemish Agency for Persons with Disabilities (VAPH), Flemish Social Protection Agency, Flemish Agency Uitbetaling Groeipakket, Agency Opgroeien, Agence pour une Vie de Qualité (AVIQ), Iriscare, Personne Handicapée Autonomie Recherchée (PHARE), Dienststelle für Selbstbestimmtes Leben (DSL), Fedris, RIZIV, and FPS Finance. Many other (regional) bodies are also likely to keep relevant data, including employment services, educational institutions, and mental health providers. These sources of administrative data will presumably also be mapped by the Working Group on Statistics in the future.

[73] Sometimes these data are even scattered within a region (e.g., in Flanders, data on minors are mainly with the Agency Opgroeien rather than with the VAPH).

[74] See, for example, Goedemé et al (2022). TAKE: Reducing poverty through improving the take up of social policies. Final Report.

applications, on persons on a waiting list, or on the nature of the disability (for example, the Directorate-General for Persons with Disabilities registers the specific pathology, but this registration is optional and limited to one 'main pathology'[75]).

- Finally, questions can be raised about the **objectivity** of the data. Both on the side of the person applying and on the side of the agency, certain elements may result in a biased assessment of disability. The persons being assessed may be afraid of missing out on the right, or conversely, may be motivated to show what they are capable of. At the same time, the agency may be under pressure to estimate the person's needs as lower than they are, due to a lack of funding.

ADMINISTRATIVE DATA: TAKE-AWAY

A wealth of administrative data exists in Belgium that can serve as a proxy for the presence of disability and related concepts.

However, the complex distribution of jurisdiction, the non-take-up of rights, the limited comparability of the data, and the lack of transparent documentation pose major challenges to their use. Moreover, claims that they are more objective than other data should be treated with caution.

4.1.4. Other proxy variables

Our crowdsourcing exercise revealed that disability and other variables are operationalised in a number of **additional ways** in the collection of (equality) data. For instance, people are sometimes approached for research or data collection because they are **part of a certain population**: because they live in a facility for people with intellectual disabilities, because they attend(ed) special education, because they are a member of a patients' association, because they use sign language, ... In some cases, this is because the researchers are specifically interested in the population in question. In other cases belonging to the population is treated as a proxy for (a specific) disability. **Self-selection** is also regularly used: if a person participates in a study, they will probably belong to the target group.

Such creative ways of operationalising disability and other concepts can certainly be appropriate or even essential in some cases, especially when one wants to survey certain target groups or gain insight into a very specific (sub)domain. However, it often involves a type of external definition and categorisation. Moreover, the data will often not be comparable with other data because a very specific operationalisation is used.

USING OTHER PROXIES: TAKE-AWAY

We consider it important to be careful when creating proxy variables yourself, as this often involves an **external definition** that the person themselves may not agree with.

Where possible, it would be good to additionally ask one of the questions in 4.1.2, ensuring that the person can define **themselves** and that the data also contain a more **standard measurement**.

[75] Iweps (2019). Les Wallons reconnus en situation de handicap : perspective statistique, p. 16.

4.1.5. Self-identification of disability

We have repeatedly referred to **self-identification** as a good practice according to the human rights-based approach to equality data. The IEDCB I report also discussed the relationship between external definition and self-identification and made the following recommendation: **Whenever possible and appropriate: strive to use self-identification**[76].

To speak of full-fledged self-identification, the data collection must meet a number of conditions[77]:

- The individuals **themselves** provide the data on their personal characteristics.
- They can **refuse** to do so and a clear option should be provided for this.
- They can choose **how** to answer the question. There should be the option to fill in an open answer and to indicate multiple categories. The second point is perhaps less crucial for the current discrimination grounds than for grounds such as ethnic origin, gender identity, or sexual orientation. However, we do see that existing data do not always capture the presence of multiple disabilities or conditions, which is problematic.
- It is strongly recommended that the questions and categories be developed in a **participatory manner**, so that the individuals concerned by the data have a say in how the data are collected and can respond to the questions in an optimal way.

To what extent is **self-identification** already used in equality data collection related to disability? What are the views of the different stakeholders in this regard?

The **crowdsourcing** showed that many actors already work with self-identification of disability to some extent and are open to using it. At the same time, people indicated that information gathered through self-identification is often complemented by other sources of information (e.g., when granting reasonable accommodations in higher education, self-reported disability is taken into account, but also medical information). Moreover, in certain situations, self-identification is not possible at all (e.g., for persons with severe intellectual disabilities).

While administrative data are not usually collected via self-identification, this is usually an option in surveys or qualitative research. However, if we scrutinise the **disability questions** covered in section 4.1.2, we see that they do not necessarily meet each of the above conditions.

In terms of the **individuals themselves** providing the data, it is questionable whether someone really **identifies** themselves as a person with a disability via instruments such as the GALI or the WG-SS. Often, 'disability' is not mentioned literally in the question; rather, the question contains wording or an underlying logic that matches the UN definition. In other words, those who *draft* the question define when someone has a disability. Moreover, answers are often further **recoded** after data collection in order to determine who has a disability and who does not (especially in the case of the WG-SS). Therefore, the final categorisation may differ from how a person sees themselves. The crowdsourcing cited **alternatives** that rely more clearly on self-identification, such as "Do you consider yourself a person with a disability?". These represent an interesting option but, to our knowledge, are not yet widespread and validated. Finally, questions about disability are sometimes answered by a **person in the environment** rather than by the person themselves. While this is of course unavoidable in some

[76] Unia (2021). Final report 'Improving equality data collection in Belgium', p. 31.
[77] European Commission (2016). European Handbook on Equality Data - 2016 Revision.
Subgroup on Equality Data (2018). Guidelines on improving the collection and use of equality data.
Unia (2021). Final report 'Improving equality data collection in Belgium'.

cases, there is a clear need to reflect on ways to adapt questions and methods in order to maximise the likelihood that persons are able to answer themselves.

The **response options** also do not consistently meet the conditions for self-identification. For instance, it is not always clear to respondents that they can **refuse** to answer the question. The possibility of recording a refusal to answer is mentioned in the instructions for interviewers, but this is usually not explicitly provided as an option for the respondent[78]. It is also usually not possible to give an **open answer** (e.g., to mention an impairment not included in the WG-SS). Once again, an interesting alternative was mentioned in the crowdsourcing: "Can you describe your situation in a few words?" This is obviously less suitable for large-scale data collection because it greatly complicates data analysis.

Finally, the *WG/UNICEF Child Functioning Module* is a good example of a tool that was developed in a **participatory manner**: among others, organisations representing persons with disabilities were involved in its development and validation[79]. For most of the other instruments, however, it is not clear whether and to what extent persons with disabilities were consulted.

In sum, there is certainly **room for improvement** to better align data collection based on self-identification with the human rights-based approach.

To conclude this section, we briefly reflect on the **advantages and disadvantages of self-identification**.

We can distinguish the following advantages of self-identification of disability:

- Data are collected with **respect** for the human rights of the individual.
- A **situational definition** of disability can be used, by asking about the respondent's perception and experience of environmental barriers.
- As responses can change with repeated data collection, self-identification takes into account the **dynamic nature** of disability.
- When the person can answer anonymously, self-identification can help to measure **invisible limitations**, including those that are **sensitive** (e.g., limitations in sexual functioning).

However, the crowdsourcing also highlighted several disadvantages of self-identification:

- Additional research is needed to verify the **validity and reliability** of new questions and accurately assess the risk of **over- and under-reporting**.
- Self-identification says nothing about which rights a person is actually entitled to and what services they can access. Sometimes it is more relevant to know what **specific needs** someone has than to know how they sees themselves.
- **Analysing** the answers to self-identification questions can be challenging**,** especially when there are many open-ended responses or when many people refuse to answer.
- Using self-identification is less evident for **specific diagnoses or conditions** for which there are established definitions (e.g., diabetes depends on certain thresholds).
- **Discrimination** is often based on how others perceive the person. How someone perceives themselves is therefore not necessarily linked to whether they experience discrimination. If

[78] Eurostat (2020). Standardised key social variables: Implementing guidelines, p. 67.
Washington Group on Disability Statistics (2023). Interviewer guidelines, p. 6.
[79] UNICEF (2021). Seen, Counted, Included: Using data to shed light on the well-being of children with disabilities.

the focus lies on discrimination, it may be more relevant to (also) include a question that measures **auto-hetero perception**, where people indicate how they are perceived by others[80].

SELF-IDENTIFICATION: TAKE-AWAY

We do not claim that self-identification is always the best option nor that it is suitable for every purpose. However, we do recommend to always consider self-identification (possibly as an additional source of information) and to adhere to the conditions for full-fledged self-identification as much as possible.

4.1.6. Conclusions

We have offered here an overview of the ways in which disability is currently operationalised in equality data and the aspects that can still be improved. For those who wish to compile equality data - either by using existing data or by collecting new data - we hope that this overview helps them to reflect critically on how to measure disability.

We do not recommend one specific question or method of operationalisation. **Which method is most appropriate depends on the type of equality data one wants to generate.** In some cases, it may be relevant to know whether someone has a disability in general terms, while in other cases it may be more appropriate to measure specific impairments or needs. Sometimes it may be necessary to work with a **pragmatic definition** or an available proxy of disability, especially when the data are collected to inform a specific policy. Finally, it can be a good idea to **combine** multiple sources of information, so that they complement each other and partly compensate for the individual limitations of each. Here, of course, it is important to be transparent about which operationalisation was used for each of the reported results[81].

4.2. Thematic focus: avoiding confusion between weight discrimination and discrimination against people with obesity

DISCLAIMER: In this section, we use the word 'fat'. We realise that this word has a negative connotation for many. However, organisations fighting against weight stigma and discrimination want to reclaim the word 'fat' as a neutral way of describing a person[82]. We adopt the same word choice here and thus use 'fat' as a neutral term.

Although we generally received little input regarding the discrimination ground of **physical characteristics**, stakeholders did draw our attention to a phenomenon related to it, namely **weight stigma and discrimination** (sometimes also called *fat phobia* in English; *grossophobie* in French, *vetfobie* or *dikfobie* in Dutch[83]). All these terms refer to discrimination and stigmatisation of fat people

[80] For more on auto-hetero perception, see Unia (2021). Final report 'Improving equality data collection in Belgium'.

[81] For an example of a recent report combining a large number of data sources, see Statistics Flanders (2022). Maatschappelijke positie en participatie van personen met een handicap.

[82] Website of fat friendly
Website of Dik Voor Mekaar

[83] Note that weight discrimination and fat phobia are technically not synonymous. Fat phobia explicitly refers to discrimination against fat people, which is not entirely equivalent to discrimination based on weight. For example, a person can be very muscular and therefore weigh a lot, while not looking fat and therefore not experiencing this form of discrimination. In general, however, weight discrimination is most commonly used to refer to discrimination of fat people.

(i.e., on the basis of their weight and/or body size)[84]. Our inventory shows that more equality data are needed on weight discrimination in Belgium.

A challenge in collecting and interpreting such data is how relevant terms are used and defined. Indeed, some stakeholders working on the topic of weight discrimination advocate for systematically **avoiding** terms such as 'overweight' and 'obesity' because they medicalise being fat. Other stakeholders emphasise the importance of defining obesity as a **chronic disease**. At first glance, these positions seem to conflict with each other, especially in terms of what terms are being used. We therefore noticed that this was sometimes a sensitive issue in the advisory group. However, if we analyse both positions more thoroughly, the contradictions immediately seem less pronounced.

In this section, we want to outline the two different perspectives and illustrate how the same terms are sometimes used and interpreted in different ways. In doing so, we want to encourage **reflection and dialogue on the terms that are used and their definitions**. This is crucial to avoid confusion between different issues, to avoid contributing (unwittingly) to stigmatisation, and to improve the quality of data on these types of discrimination.

4.2.1. What is weight discrimination?

Weight discrimination refers to the phenomenon whereby fat people are discriminated against or stigmatised. Several examples were cited during the crowdsourcing. Weight discrimination can manifest itself in **insults or other negative reactions** towards fat people. In addition, places and services are often **not very well adapted** (e.g., public transport, aeroplanes, public toilets, sports facilities and other leisure services, etc.). **Systematic discrimination in the medical sphere** was also mentioned: fat people are often treated less respectfully by health professionals, doctors' heavy focus on weight can lead to wrong or delayed diagnoses, medical equipment is not always adapted which can lead to wrong measurements, certain treatments are less effective (e.g., the morning-after pill), and so on.

Which **protected ground** applies in the case of weight discrimination depends on the situation, as the above examples show. Sometimes it appears to involve discrimination based on **physical characteristics** (i.e. how one looks), which stems from the fact that there is a negative image of fat people in society and/or immediate assumptions are made about their health. Other cases involve an environment being poorly adapted (e.g., seats being too small), which results in participation problems. In such cases, the ground of **disability** seems to apply (see definition in section 1.3.1).

Stakeholders stressed the importance of **not viewing weight discrimination as discrimination based on health status**. This would imply that being fat equals having a disease or being unhealthy. This is also why it is advocated **not to use the terms 'overweight' and 'obese'**. Stakeholders explained that these terms do not constitute a neutral description of a person and that they medicalise being fat.

As we discuss in more detail below, the World Health Organisation defines obesity as a **complex chronic disease** characterised by abnormal and/or excessive fat deposits that can impair health[85]. This definition does not equate the disease obesity per se with how much one weighs or how fat one is. However, if we look at the concrete definition of the categories 'overweight' and 'obese', we see that this is done on the basis of a proxy variable, namely the **Body Mass Index (BMI):** from a BMI of 25

---

[84] Puhl, R. (2021). Weight stigma is a burden around the world – and has negative consequences everywhere. *The Conversation.*

[85] World Health Organization (2019). International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision). Code 5B81: obesity.

European Commission (2023). Health Promotion and Disease Prevention Knowledge Gateway: obesity prevention.

onwards, one is said to be overweight, from a BMI of 30 onwards one is said to be obese (for adults)[86]. Thus, although BMI merely reflects the relationship between height and weight, it is used to operationalise chronic disease. **In this way, being fat is thus indeed medicalised**.

We conclude that it is **not appropriate** to use the term **obesity** in the context of weight discrimination, as this term refers to a chronic disease and therefore **does not** constitute **a synonym for 'being fat'**. **To understand and recognise weight discrimination, individuals' weight or appearance must be considered separately from their health status**.

4.2.2. Discrimination against persons with obesity

As we mentioned above, the World Health Organisation classifies **obesity** as a **chronic complex disease** defined by abnormal and/or excessive adiposity that can impair health. In other words, obesity is characterised by **problems with the functioning of adipose tissue**, its amount and/or its distribution[87].

Although obesity is included in the *International Statistical Classification of Diseases and Related Health Problems* (ICD), stakeholders working on this topic indicated that there is a lot of **misinformation and confusion surrounding this disease**, including in the medical community.

For instance, obesity is often equated with how much someone weighs or how fat someone is. However, if we look at the definition above, obesity is essentially not about weight and can be separate from it. Although weight can be one visible sign of the disease (especially in the absence of correct treatment), and although obesity is more often diagnosed in fat people, the above-mentioned problems with the functioning of adipose tissue can also occur in people who are not fat. It is therefore far from ideal to use weight or BMI as the sole indicator of obesity, as it does not adequately reflect the complexity of the disease[88].

The result of this misinformation and confusion is that obesity is not sufficiently seen as a chronic disease and that attention is not always given to the right things. According to stakeholders, healthcare workers often focus heavily on weight even though it does not necessarily reflect a person's health status. As a result, patients with the disease of obesity often struggle to obtain a **correct diagnosis and appropriate treatment**, and their right to well-informed care is compromised. Failure to treat people with obesity correctly amounts to discrimination based on the ground of **health status** (or possibly the ground of disability, given the chronic nature of the disease and the possibility of participation problems).

The stakeholders mentioned that a lexicon for obesity is currently being developed[89], with the aim of clarifying the term 'obesity' (and other related terms) and arriving at a shared language. We conclude that it is crucial to **use these terms correctly** and to consider obesity as a disease that can be independent of a person’s weight or body size.

---

[86] World Health Organization (2024). Obesity and overweight.

[87] Bowman-Busato, J. et al (submitted for publication). Providing a common language for obesity: The European Association for the Study of Obesity obesity taxonomy.

[88] EASO (2023). Pre-Read for 26 October 2023 Technical Working Session on Overcoming Policy Barriers to implementing obesity under the NCD Framework.
Frühbeck et al (2019). The ABCD of Obesity: An EASO Position Statement on a Diagnostic Term with Clinical and Scientific Implications. *Obesity Facts, 12*(2), p. 131-136.
Bowman-Busato, J. et al (submitted for publication). Providing a common language for obesity: The European Association for the Study of Obesity obesity taxonomy.

[89] Bowman-Busato, J. et al (submitted for publication). Providing a common language for obesity: The European Association for the Study of Obesity obesity taxonomy.

4.2.3. Conclusions

The above shows that the two perspectives do not necessarily have to contradict each other. Rather, they seem to represent two sides of the same coin. Those who want to create awareness about **weight discrimination** find that **assumptions about health** often divert attention away from the issue. Those who advocate for better diagnosis and treatment of **obesity** find that **assumptions linked to weight and appearance** often distract from this. Both views amount to saying that health, weight, and appearance should all be considered separately.

When collecting **equality data on these forms of discrimination**, it is therefore essential to always make clear what one is talking about and to explicitly define the terms that are used. This is important not only to increase the comparability of the data, but also to avoid the data contributing to further stigmatisation.

WEIGHT DISCRIMINATION AND THE USE OF THE TERM OBESITY: TAKE-AWAY

Body size, weight, and health are different concepts and can lead to **different forms of discrimination**. Therefore, avoid language that **confounds** them.

Are you making **assumptions** about the link between weight and health? If so, state them explicitly.

Weight discrimination refers to discrimination against fat people regardless of their health status. In this context, consider using the word 'fat' as a neutral way of describing people, possibly accompanied by additional explanations to remove the negative connotation of the word.

Obesity and overweight are terms with a medical meaning. Do **not** use them as synonymous with being fat. Use the term 'obesity' to refer to the chronic disease characterised by problems with adipose tissue, regardless of one's weight or body size. If you operationalise obesity solely through BMI, state this explicitly and note the necessary caveats.

5. Recommendations

Based on the results discussed in the previous chapters, in this final chapter we formulate a number of **concrete recommendations**. We focus mainly on how Belgian equality data related to the discrimination grounds of disability, health status, and physical characteristics can be improved in the future, although recommendations 5 to 7 also apply to equality data in general.

Our recommendations are largely addressed to **actors who regularly collect and/or analyse data**: statistical institutions, knowledge centres, academic researchers, administrations or civil society organisations planning to conduct a study on a specific topic, and so on.

However, it is important to emphasise that our recommendations are also addressed to **political actors and policymakers** (e.g., cabinets and administrations). After all, they have to create the necessary conditions for the above actors to effectively collect and disseminate high-quality equality data. For this, political support must be established, resources and time must be provided, cooperation between different actors and levels must be organised and maintained, and so on. Moreover, these actors bear the important responsibility of actually using equality data in their work (see recommendation 7).

5.1. Fill in the gaps

In Chapter 3, we highlighted several gaps: **topics about which there is still too little or no equality data available, but which stakeholders consider crucial**. Therefore, a logical first recommendation is to **fill these gaps**.

1. **Collect disability-related data in a more systematic manner, so that the impact of policies on persons with disabilities (and any resulting inequalities) can be effectively monitored**. A 'twin-track approach' is crucial here[90]. This means that, on the one hand, the **impact of disability-specific policies** and the operation of services/systems specifically intended for persons with disabilities should be monitored. This requires accurate and complete data on these systems (e.g., data on how often certain rights are granted, as well as any associated waiting lists). On the other hand, the **impact of mainstream policies on persons with disabilities** must also be monitored in different areas of policymaking. For this, data in general should more routinely allow to distinguish between persons with and without disabilities. This therefore requires some form of *disability mainstreaming*[91] in data collection: including disability as a variable by default in data collection efforts, where appropriate and relevant.
2. **Collect more data on inequality and discrimination based on disability in the domains of justice and police, media, and digitalisation**. The gap analysis showed that there is still little equality data related to disability in these domains. However, it seems highly important to have adequate data on these topics. What is the situation regarding access to justice and the legal capacity of persons with disabilities? How are persons with (psychosocial) disabilities treated by the police? To what extent are the rights of interned persons being respected and are they

[90] Committee on the Rights of Persons with Disabilities (2016). Rules of procedure, Annex: Guidelines on independent monitoring frameworks and their participation in the work of the Committee on the Rights of Persons with Disabilities, pp. 33-43.
European Union Agency for Fundamental Rights (2023). Implementing the UN Convention on the Rights of Persons with Disabilities: Human rights indicators.
[91] *Disability mainstreaming* refers to the principle that the effect on persons with disabilities should by default be considered when developing any policy. See Unia. Handistreaming: toolbox.

able to proceed into the regular care system?[92] How often do persons with disabilities appear in the media and how are they represented? How does the increasing digitalisation affect access to information and rights?

3. **Collect the necessary data to monitor progress towards a truly inclusive society.** Persons with disabilities are still too often forced to rely on separate systems (e.g., special education, companies for adapted work, specific living facilities) because society as a whole is insufficiently adapted to their situation. To monitor the evolution towards an inclusive society, there is a need for reliable and repeated data on the situation of persons with disabilities. Also collect information on the situational aspects (barriers in the environment) that prevent persons with disabilities from fully participating in society. An important example would be high-quality data on the accessibility of buildings, websites, etc. for persons with disabilities.
4. **Thoroughly map the problem of non-access to and non-take-up of rights by persons with disabilities.** Data on non-take-up of rights are in themselves a valuable source of equality data, as they indicate who does or does not have access to their rights. In addition, they also constitute important contextual information for those who wish to compile equality data based on administrative data, as they allow to estimate who is (not) captured by each data source.
5. **Collect data on discrimination based on health status, separately from disability, and keep these discrimination grounds clearly separate in data collection.** We have come across relatively few data sources that can demonstrate inequality or discrimination based on the ground of health status, without it being confounded with the ground of disability. There is therefore a need to keep these two grounds clearly separate in data collection and to possibly collect new data specifically aimed at measuring discrimination based on health status. Moreover, the analysis of the inventory suggests that there is a lack of data on inequality based on health status in the domains of education and poverty. Thus, additional equality data on health status should be collected in these domains.
6. **Reflect on intersectional analyses before collecting data**.
    - Existing data mainly allow us to look at the intersection with gender on the one hand and age on the other. This is good as the UN Convention on the Rights of Persons with Disabilities calls for specific attention to the **situation of children and women with disabilities**[93]. However, being able to cross variables is not always sufficient to gain insight into how personal characteristics interact and lead to specific situations of inequality. There is therefore a need for **full-fledged intersectional analyses**, for instance in the form of qualitative research.
    - There are far less data available that contain information on the intersection with other characteristics. Ensure that data - where relevant - allow to look at the **intersection with social origin/situation, national/ethnic origin, skin colour, language, family situation, or sexual orientation**. Of course, it is not relevant to take all of these characteristics into account and analyse each intersection every time (moreover, this requires very large samples). However, it is important to consider **from the start of data collection** which other characteristics could in combination lead to particularly disadvantaged situations.
    - Also consider the necessary sample size and the minimum number of individuals per intersection required to actually conduct the intersectional analyses. If necessary, work

[92] For a recent overview of challenges linked to the Belgian system of internment, see Unia (2023). Réinsertion des personnes internées : quels défis dans un État de droit ?

[93] United Nations (2006). Convention on the Rights of Persons with Disabilities, art. 1.
Committee on the Rights of Persons with Disabilities (2014). Concluding observations on the initial report of Belgium.

with **targeted sampling** to ensure that enough individuals are reached at each intersection.

7. **Consider including disability, health status, and/or physical characteristics when conducting discrimination tests**. The analysis of the inventory (which admittedly does not include purely academic sources) suggests that little discrimination testing happens for these grounds. However, this is crucial to obtain strong, experimental evidence of discrimination and to uncover any invisible discrimination in various domains. So take these grounds into consideration when planning discrimination tests, especially in domains where there is a strong suspicion of discrimination based on these grounds.
8. **Where possible, make local data available (e.g., at the level of the municipality)**. It will not always be possible to make these data publicly available by default, especially for smaller populations, as the privacy of individuals cannot then be guaranteed. In any case, however, it would be good to always clearly describe what data are available and to indicate whether and how they can be requested for data analysis in a secure environment (see also section 5.6).

5.2. Reflect on the most appropriate operationalisation of disability

In Chapter 4, we described how disability is typically measured when collecting equality data. The many pitfalls and concerns we raised illustrate the need for careful reflection when operationalising this concept.

1. **Use the definition from the UN Convention** (rather than a medical or administrative definition) so that no one is overlooked.
2. **Avoid mixing the concepts of disability and health status** within the same variable or question.
3. **Choose the method that is most appropriate for the question at hand** (see section 4.1). Consider which concepts should definitely be measured (e.g., simply the presence of a disability, or also its severity and nature). In addition, each method has advantages and disadvantages. Which advantages are decisive in the context of the specific objective? Which disadvantages would prevent the objective from actually being achieved?
4. **Where possible and appropriate, ensure that the data also include information on the nature of the disability.** It goes without saying that some subgroups of persons with disabilities (e.g., people who are deaf or hard of hearing) face partially different barriers than other subgroups (e.g., persons with reduced mobility). As such, inequalities do not necessarily manifest themselves in the same way for all persons with disabilities. Data that can be disaggregated according to the nature of the disability are needed in order to identify who is particularly disadvantaged in a specific domain or by a specific policy measure. This can be done, for example, by adding the *Washington Group Short Set* to a survey.
5. **Preferably use an existing method of operationalisation**, if one already exists that meets the objectives of the data collection. Stay as close as possible to the original method in terms of wording and implementation.
6. **Organise consultation between different statistical and research institutions so that measurement instruments are systematically implemented and translated in the same way**. Existing surveys often contain similar instruments, but with subtle differences in terms of the exact wording, translation, or implementation. In addition, there is not yet an official or established Dutch version of the *Washington Group Short Set*. To address these issues, consultation between different actors is crucial.
7. **Keep in mind that administrative data inherently involve a proxy of disability and can be distorted by several factors**, such as differences in the conditions of recognition, exclusion criteria unrelated to disability, non-take-up of the right in question, etc.

8. **Consider self-identification (see section 4.1.5) as a possible (perhaps complementary) method.** In any case, apply its main principles as much as possible: provide a clear option to refuse, involve the target group in developing the questions, etc.
9. **Consider asking directly about discrimination or using auto-hetero perception.** If the aim is not so much to detect situations of general inequality, but rather to measure discrimination specifically, it may be more relevant to ask directly about experiences of discrimination on the basis of disability (as is done, for example, in the EU-MIDIS II[94]). Another interesting option is to use an auto-hetero perception question, in order to disaggregate the data based on how people think they are perceived by others.

5.3. Collect data on discrimination based on physical characteristics

Both in our inventory and in our crowdsourcing, we noticed that the **ground of physical characteristics** came up very little, even though we specifically tried to reach and involve actors working on this topic. Only for the phenomenon of weight discrimination (see section 4.2.1) were we able to gather more information during the project, and even with regard to this topic there still appears to be a lot of confusion and a lack of concrete (Belgian) equality data.

Therefore, our third recommendation is to **map discrimination based on physical characteristics in Belgium.** This requires at least the following steps:

1. **Establish a definition** that remains sufficiently close to the legally protected discrimination ground, while also being adapted to sociological realities.
2. **Develop and validate instruments** that allow this ground (or discrimination based on this ground) to be operationalised in the context of data collection. For this ground, a particularly relevant avenue seems to us to develop questions that directly probe discrimination based on this ground and/or questions that rely on auto-hetero perception.
3. **Work out a coordinated approach** to collect the necessary equality data, both through integrating this ground into existing forms of data collection (e.g., surveys on discrimination) and through setting up new initiatives that focus specifically on this topic (e.g., qualitative research, field tests).

5.4. Adapt research methods to better reach specific target groups

The crowdsourcing revealed that **certain groups** are insufficiently or not at all represented in the available data. This may be because certain individuals are not part of the sampling frame for large-scale surveys, because persons with a rare condition are statistically unlikely to be well represented, or because some persons are not surveyable or very difficult to survey when standard methods are used.

Especially for surveys that are the source of important equality data regarding disability, it is crucial to ensure that not only the subset of people who do not experience barriers to participating in research is reached. In that case, certain groups will remain invisible and the true inequality is likely to be **underestimated**. Depending on the objective of the data collection, the above issues need to be tackled to achieve high-quality and complete equality data, where no relevant group remains invisible.

1. **Reflect on how to better reach persons in facilities for collective living (institutions)**. The way in which the sampling frame for surveys is constructed usually focuses on reaching persons in 'traditional' households, which excludes a priori the population of persons living in institutions.

[94] European Union Agency for Fundamental Rights (2017). Second European Union Minorities and Discrimination Survey - Technical report.

Consider other methods of sampling that do reach this population. Consult with key actors (e.g., facility staff) to understand important prerequisites for organising repeated and standardised data collection in this context.

2. **Consider targeted sampling (also called purposive sampling or using booster samples*)* when it is important to reach a sufficient number of individuals from specific groups** and when this is unlikely to happen through random sampling. Here, individuals from the target groups are specifically approached so that they are 'over-represented'[95]. While this will result in the sample no longer being representative of the entire population, the data will allow statements to be made about individuals from the target groups. When this technique is used, it is important to state this explicitly so that the user of the data can assess its representativeness.
3. **If necessary, adapt the methods to maximise the chances that persons with disabilities can participate in research themselves**.
    - In some cases, this is not possible and it will remain necessary to collect data through a person in the environment, as in the case of persons with severe intellectual disabilities or in the case of young children with disabilities (though it is still advisable to use appropriate tools, such as the *WG/UNICEF Child Functioning Module*).
    - In other cases, procedures can be provided to quickly and easily request reasonable accommodations (e.g., working with a sign language interpreter)[96].
    - Finally, the method can be adapted to be more accessible to different groups by default, for example by providing an easy-to-read version, using accessible forms and audio description, providing a translation to sign language by default (via videos), and so on[97].

5.5. Create the necessary conditions for true participation of concerned groups in the process of data collection about them

"Nothing about us without us". When data collection is almost entirely in the hands of people who are not part of the target group, which is de facto still very often the case, the people involved have **no say** in a process that specifically concerns them. This is problematic and goes against the human rights-based approach to data.

Therefore, we would once again like to strongly emphasise the **importance of participation in the collection of equality data**. Based on the crowdsourcing and our own experiences during this project, we found that this principle is often still insufficiently applied.

For instance, target groups are often not consulted during the **entire data collection process**. Usually, for example, they are not contacted until the research question and methods are already largely established, leaving little room to integrate their advice. It is crucial to involve the people concerned from the very first steps of the process, so that they can help determine on which topic data should be collected and in what way.

We also received signals from civil society organisations that they are often asked to fulfil an advisory role in research projects, but that this role is not always taken seriously. It sometimes seems to come down to a **participation *pro forma*,** with their advice not actually being acted upon. Moreover, the

[95] European Commission (2016). European Handbook on Equality Data - 2016 Revision, p. 40-41.
[96] UNICEF (2020). Producing disability-inclusive data: Why it matters and what it takes.
[97] See, for example, Fontaine, S. (2015). Enquêter auprès des sourds. Implications éthiques, méthodologiques et statistiques de l'adaptation d'enquêtes d'opinion au public sourd, for a comprehensive overview of how surveys can be adapted to the situation of persons who are deaf or hard of hearing.

request for advice often comes very late in the process and hence no real adjustments are possible anymore. This is particularly unfortunate as valuable expertise is not utilised, whereas the organisations do spend valuable time and resources to provide it.

So while a participatory approach is increasingly becoming a part of research projects, the time and resources provided are not always adapted to this. **However, a participatory process absolutely does require additional time and resources**: efforts need to be made to reach individuals, explain the research to them, and involve them on a regular basis; their advice must be taken into account, which requires sufficient flexibility in methods and timing; and conditions need to be created to ensure that everyone can participate on an equal basis, which includes providing reasonable accommodations, reimbursing travel expenses if necessary, or considering possible compensation for those who sacrifice their free time to participate (which may be the case for advocacy groups or volunteer organisations).

Thus, in order to make participation an integral part of equality data collection, a **shift in mindset** still seems required. Value the involvement and advice of the groups concerned and create conditions that allow their full participation, including the provision of sufficient time, resources, and flexibility.

5.6. Make and keep equality data available and accessible

There are a lot of equality data, but also several barriers that often prevent them from being easily found or used. With a view to ensuring that equality data are and remain accessible, we recommend the following:

1. **Ensure transparency and sufficiently comprehensive documentation**. Especially in the case of administrative data, but also in the case of some surveys, it is particularly difficult from someone on the outside to gain insight into what data are collected or kept, how this is done, in what form they could be made available for data analysis in a secure environment, and how they could possibly be linked to other data. This can be remedied by making information on all of these aspects publicly available.
2. **Make sources of equality data accessible to persons with disabilities.** For example, ensure that reports and websites comply with digital accessibility standards[98]. Also consider making available a version (or a summary) in easy-to-read and/or in sign language.
3. **Make sure that reports remain findable in the long term.** While updating the sources in the data hub, we stumbled upon a seemingly trivial problem, but one that actually has a major impact: a large number of links to reports or web pages were broken. In some cases, the source could no longer be found at all (including some recent sources). Sometimes this was due to a restructuring of the website in question; in other cases a report had been linked to a news item, for example, but not been saved anywhere else. This is particularly unfortunate as these are data sources into which a lot of work, time, and resources has gone. In order to make maximum use of equality data, it is crucial that they remain available in the long term. Efforts should therefore be made to create central databases of publications, in which sources are stored in a systematic and permanent way. Consideration could also be given to the use of persistent identifiers, which allow sources to be retrieved easily even if the structure of a website has changed[99].

[98] See, for example, the AnySurfer website. Standards et lois.
[99] An example is the Digital Object Identifier (DOI). For more comprehensive information on persistent identifiers and preserving digital material, see Digital Preservation Coalition (2015). Digital Preservation Handbook, 2nd Edition.

5.7. Make maximum use of equality data

The aim of Unia's data hub is to make it easy to find equality data for specific discrimination grounds, as well as for specific domains and levels of government. In the coming months and years, we will continue to work on making this tool known to potential users and on improving its user-friendliness. **This is an important first step, but the responsibility of actually working with equality data lies with the user.**

We therefore strongly recommend making more use of the wealth of useful information that is contained in equality data.

In the introduction, we listed various purposes for which equality data can be particularly valuable: monitoring inequality and discrimination, making discrimination visible, understanding the structural mechanisms that contribute to inequality and discrimination, designing and evaluating policies based on facts rather than intuition, raising awareness about discrimination, strengthening evidence in legal cases, and reporting on human rights treaties (such as the UN Convention on the Rights of Persons with Disabilities). For some of these purposes we find that equality data are already being used, for others we still see a significant under-use of equality data.

1. **Consider how inequality and discrimination can be monitored more systematically in Belgium**. Equality bodies and human rights institutions (including Unia) could, for instance, publish **reports** on a periodic basis, combining equality data from different types of sources to monitor the state of equality and identify concrete areas for action. Specifically for the ground of disability, some reports already exist that partially do this, which may serve as inspiration[100]. For reports not limited to one ground, there are also some interesting international examples[101]. Besides quantitative data, it is important that such reports also draw on qualitative research and other forms of data that can provide insight into the **underlying, structural mechanisms**.
2. **Use relevant equality data when developing new policy measures** so that they are based on the most accurate picture possible of reality. Also immediately consider what role (those same) sources of equality data can play in **evaluating** the impact of policies and actually use them to do so. For example, explicitly state in policy statements on what data and sources the proposed actions are based, and then in subsequent years revisit those same data and sources in the policy statement (if updates are available), in order to assess evolutions over time.
3. **Reflect on how equality data can be used as (supporting) evidence in legal proceedings**, what form they should take to be appropriate for this, and how the data hub can better respond to lawyers' needs regarding equality data.
4. **Look into how equality data can be used even more systematically in the context of the UN Convention on the Rights of Persons with Disabilities**, both in terms of reporting by the Belgian state and in terms of independent monitoring by organisations (including Unia). Although equality data are regularly used for this purpose, this is still done in a rather unsystematic way. We refer in this regard to the work of the *Fundamental Rights Agency,* which has developed concrete guidelines for applying the **OHCHR approach to human rights**

[100] GRIP vzw (2016). Inclusiespiegel Vlaanderen 2016.
Statistics Flanders (2022). Maatschappelijke positie en participatie van personen met een handicap.
[101] The Equality Ombudsman (2023). The state of discrimination 2023. Annual report from the Equality Ombudsman.
The Centre for Human Rights Resources and the Office for Non-Discrimination, Barcelona City Council (2023). Barcelona Discrimination Observatory Report 2022.

indicators[102] to the monitoring of the UN Convention[103]. This approach distinguishes **three types of indicators**: **structural indicators** that reflect government commitment (e.g., legislation, adoption of policies, provision of complaints mechanisms); **process indicators** that indicate effort (e.g., allocation of funds, implementation of policies, effectiveness of complaints mechanisms); and **outcome indicators** that reflect actual results (e.g., impact of policies and other measures, prevalence of discrimination, factual inequality). Note that the equality data we centralize in the data hub are particularly useful to feed the latter category of indicators: they provide insight into the concrete results of policies for the **day-to-day reality of persons with disabilities**. In sum, we recommend applying this approach whenever possible in the context of the UN Convention[104].

[102] Office of the High Commissioner for Human Rights (2012). Human Rights Indicators: A Guide for Measurement and Implementation.

[103] European Union Agency for Fundamental Rights (2023). Implementing the UN Convention on the Rights of Persons with Disabilities: Human rights indicators.

[104] One challenge in this regard is how to determine the indicators. For some articles of the UN Convention, indicators have already been developed by the European Union Agency for Fundamental Rights, such as for Article 19 (Human rights indicators on Article 19 CRPD) and Article 29 (The right to political participation for persons with disabilities: human rights indicators). In addition, there are many international examples from which inspiration can be drawn, see European Union Agency for Fundamental Rights (2023). Implementing the UN Convention on the Rights of Persons with Disabilities: Human rights indicators, Chapter 5: Monitoring CRPD implementation across the European Union - national practice examples, p. 31-34.

6. Annex: table of inventoried sources for the so-called 'racial' grounds, religious and philosophical beliefs, sexual orientation, gender identity, gender expression, and sex characteristics

Table B1. Number of sources containing data on a specific domain (by discrimination ground).

Domains	Discrimination grounds							
	Nationality	Skin colour	Other 'racial' grounds	Religious /philosophical beliefs	Sexual orientation	Gender identity	Gender expression	Sex characteristics
Work	82 (48.2%)	17 (48.6%)	128 (44.3%)	26 (38.8%)	26 (37.7%)	14 (31.8%)	8 (42.1%)	4 (40.0%)
Education	47 (27.6%)	13 (37.1%)	91 (31.5%)	26 (38.8%)	27 (39.1%)	14 (31.8%)	7 (36.8%)	4 (40.0%)
Health and well-being	45 (26.5%)	10 (28.6%)	81 (28.0%)	18 (26.9%)	24 (34.8%)	19 (43.2%)	11 (57.9%)	7 (70.0%)
Poverty and standard of living	37 (21.8%)	2 (5.7%)	53 (18.3%)	6 (9.0%)	4 (5.8%)	2 (4.5%)	0 (0.0%)	0 (0.0%)
Housing and living situation	53 (31.2%)	13 (37.1%)	80 (27.7%)	15 (22.4%)	14 (20.3%)	5 (11.4%)	2 (10.5%)	1 (10.0%)
Mobility and transportation	12 (7.1%)	7 (20.0%)	18 (6.2%)	9 (13.4%)	7 (10.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Justice and police	18 (10.6%)	14 (40.0%)	37 (12.8%)	17 (25.4%)	19 (27.5%)	13 (29.5%)	7 (36.8%)	5 (50.0%)
Media	7 (4.1%)	8 (22.9%)	23 (8.0%)	11 (16.4%)	10 (14.5%)	7 (15.9%)	3 (15.8%)	2 (20.0%)
Digitalisation	3 (1.8%)	0 (0.0%)	6 (2.1%)	0 (0.0%)	2 (2.9%)	1 (2.3%)	0 (0.0%)	0 (0.0%)
Other domain	69 (40.6%)	16 (45.7%)	107 (37.0%)	42 (62.7%)	42 (60.9%)	30 (68.2%)	12 (63.2%)	8 (80.0%)
No domain	14 (8.2%)	1 (2.9%)	18 (6.2%)	6 (9.0%)	1 (1.4%)	1 (2.3%)	0 (0.0%)	1 (10.0%)
Total number of sources	**170**	**35**	**289**	**67**	**69**	**44**	**19**	**10**

(1) The percentages between brackets should be read column by column. They reflect the number of sources containing data on a domain, relative to the number of sources containing data on the discrimination ground. For example, 27.6% of the sources related to the ground of nationality contain data on the domain of education.
(2) Because the same source can contain data on multiple domains, the sum per column is greater than the total number of sources for that ground (and greater than 100% for the percentages).
(3) For sources that cover multiple ground and domains, the fact that a ground and a domain are both covered in the source does not necessarily mean that the source actually looks at inequality based on that specific ground in that specific domain (although this is usually the case). The numbers in this table thus represent an upper bound for the actual number of sources containing data on the *intersection* of a ground and a domain.

Colophon

Improving Equality Data Collection in Belgium II: Final report
Brussels, June 2024

Publisher

Unia, Interfederal center for equal opportunities and the fight against racism and discriminations
Place Victor Horta 40 (bte 40), 1060 Brussels
T +32 (0)2 212 30 00

Writing: Unia
Final editing: Unia
Translation: Unia
Graphic concept and layout: Unia
Responsible editor: Patrick Charlier

This project was funded by the Equal Opportunities Department of the FPS Justice.

Ce rapport est aussi disponible en français.
Deze publicatie is ook verkrijgbaar in het Nederlands.

You can find this publication at www.unia.be in the 'Knowledge & Recommendations' section.

Prefer to read on paper? On the publication page of this report you can request a printed version, via the 'Order printed version' button. We will then send you a free copy; only for large quantities or an express shipment, postage may be charged.