WHODAS disability assessment pilot in four regions of Italy

Evaluation report (Output 7.3 full-pilot)

OECD

PRESIDENZA DEL CONSIGLIO DEI MINISTRI
PALAZZO CHIGI
Ufficio per le politiche in favore delle persone con disabilità

Funded by the European Union

This technical support project is funded by the Technical Support Instrument, and implemented in cooperation with the European Commission's Directorate-General for Structural Reform Support (DG REFORM)

OECD Directorate for Employment, Labour and Social Affairs &
OECD Trento Centre for Local Development

WHODAS disability assessment pilot in four regions of Italy

Evaluation report (Output 7.3)

This report describes and discusses the results from a pilot of a different disability assessment tool, the WHO Disability Assessment Schedule (WHODAS), in four regions of Italy. The analysis includes observations from 3242 individuals participating in the pilot (1327 in Lombardy, 1223 in Campania, 510 in Trentino and 182 in Sardinia). Using a statistical approach, the report assesses the performance of the 36-item WHODAS questionnaire and concludes that the tool is working well in Italy and delivering plausible, coherent, and scientifically sound and robust distributions of WHODAS scores in all four pilot regions. The report also compares the WHODAS scores of the pilot sample with the corresponding ratings for civil invalidity, as pilot participants had been assessed in both ways, and presents options on how the WHODAS questionnaire could be integrated into the current way of assessing civil invalidity in Italy.

This evaluation report is Activity 7.3 (Output 7) of a technical support project funded by the Structural Reform Support Programme of the European Commission (Contract Agreement REFORM/IM2021/009) and implemented in co-operation with the Commission's Directorate-General for Structural Reform Support (DG REFORM).

OECD

PRESIDENZA DEL CONSIGLIO DEI MINISTRI
PALAZZO CHIGI
Ufficio per le politiche in favore delle persone con disabilità

Funded by the European Union

This technical support project is funded by the Technical Support Instrument, and implemented in cooperation with the European Commission's Directorate-General for Structural Reform Support (DG REFORM)

ABOUT THE PROJECT PARTNERS

The OECD Directorate for Employment, Labour and Social Affairs (ELS) leads the OECD's work on employment, social policies, international migration and health. The Directorate oversees OECD work on interrelated policy areas that help countries boost employment and skills and improve social welfare and health. www.oecd.org/els

The OECD Trento Centre for Local Development is an integral part of the OECD and its Centre for Entrepreneurship, SMEs, Regions and Cities (CFE). The Trento Centre uses a holistic "from data to practice" approach to policies for sustainable development to offer local policy analysis, advice and capacity building activities for improved policy implementation for people, firms and places. www.trento.oecd.org

The Directorate-General for Structural Reform Support (DG REFORM) coordinates and provides tailor-made technical support to EU Member States, in cooperation with the relevant Commission services. The support is primarily provided through the Technical Support Instrument (TSI). The goal is to support Member States' efforts to design and implement resilience-enhancing reforms, thereby contributing to the EU's recovery from the COVID-19 crisis, improving the quality of public services and getting back on the path of sustainable and inclusive growth.
https://ec.europa.eu/info/departments/structural-reform-support_en

The Office for policies in support of people with disabilities is the ministerial unit in charge of coordinating the action of the Italian Government in the field of disability. In particular, its tasks relate to policies for the protection of the rights of people with disabilities and their social inclusion, in line with the UN Convention on the Rights of Persons with Disabilities and the EU Charter of Fundamental Rights. https://disabilita.governo.it/it/lufficio/

OECD Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

OECD Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on OECD Papers are welcomed, and may be sent to the OECD Trento Centre, vicolo San Marco 1, 38121 Trento, Italy or OECD Directorate for Employment, Labour and Social, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

This document, as well as any statistical data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

This document was produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

© OECD 2023

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Table of contents

1. Introduction

Disability assessment in Italy is fragmented, discretionary and outdated. The assessment of civil invalidity which determines a person's rights and entitlements to benefits and supports is limited to the identification of a medical condition, not including the considerable conceptual developments in the understanding of disability during the past three decades. Disability is now considered a social construct that is determined by a person's social, economic, and physical environment or context. This is reflected in the International Classification of Functioning, Disability and Health (short, ICF) framework, which was approved in 2001 by all WHO member states, including Italy. Ever since, the Italian government has strived to reform its system, also to comply with the UN Convention on the Rights of Persons with Disabilities which the country has ratified in 2009 – but with limited success so far. Meanwhile, pressure for reform is building up from all sides, internationally and nationally, as all stakeholders agree that reform is urgently needed. With the passage of a framework law in late 2021, Italy has made a big step towards a reform of its disability policies.

This report supports these reforms. It summarises the results of a pilot that took place in four regions of Italy (Campania, Lombardy, Sardinia, and Trentino), testing the feasibility of the inclusion of functioning information into the current assessment of civil invalidity. This is done by piloting the use of the WHO Disability Assessment Schedule (WHODAS) which was developed as a tool to identify the kind and nature of problems people are facing in their lives, in alignment with the ICF framework. The pilot is also unique insofar as the WHODAS questionnaire is implemented by social workers from the regions, not medical doctors. A great advantage of the WHODAS tool is that it has been tested successfully in many countries and different contexts; hence, Italy can draw not only on the experiences with its own pilot but also on experiences in other countries.

The report summarises findings of the pilot that took place from October 2022 to April 2023. The analyses presented in this report are based on a sample of 3242 cases from the four regions; a sample large enough to draw meaningful inferences and discuss the usefulness of the inclusion of a functioning tool into the assessment of civil invalidity. The evaluation also discusses options and makes recommendations on how functioning information, collected through the WHODAS questionnaire, can be used for, or merged into, the current way of assessing civil invalidity and how doctors (the main actors in civil invalidity assessment today) can best be supported by social workers (as implementors of the functioning tool).

This pilot is part of a larger project, conducted by the OECD together with the Italian government and funded by the European Union, on the reform of disability assessment and social protection for people with disability in Italy. The project aims to help the Italian government in identifying and overcoming the key challenges the country is facing, with the goal to support people with disability more effectively and more efficiently, and more uniformly across the country.

2. Pilot sample and WHODAS score distribution

In the ICF framework, information about categories of Activities and Participation can be collected either from the perspective of capacity (reflecting exclusively the expected ability of a person to perform activities considering their health conditions and impairments) or the perspective of performance (reflecting the actual performance of activities in the real-world environmental circumstances in which a person lives). Information about capacity typically represents the results of a clinical inference or judgment based on medical information, while performance is a true description of what occurs in a person's life. The two perspectives are therefore very different, although capacity constitutes a determinant of performance.

As the administrative act of establishing eligibility for services and supports, disability is assessed as the overall lived experience of an individual living with one or more health problems – or, in ICF terms, the level of a person's performance in light of the intrinsic health capacity and the environmental facilitators or barriers. Disability assessment is a 'whole person' assessment of the extent or level of a person's disability. This is important because a disability assessment should be a summary measure of functioning levels across different domains of actions, simple and complex, from walking, taking care of children to working at a job. Thus, the assessment must be based both on the individual's health state and specific assessments of specific activities, measured with high validity and reliability.

The ICF understands 'disability' to be any level of problem or difficulty in functioning in some domain, from the perspective of performance. The WHO has developed, tested, and recommended WHODAS as an instrument that can capture the performance of activities by an individual in his or her daily life and actual living environment. The 'actual environment' is represented in the ICF in terms of environmental factors that act either as facilitators (e.g., assistive devices, supports, home modifications) or as barriers (e.g., inaccessible houses, streets and public buildings, stigma, and discrimination). The WHODAS questionnaire, in short, is WHO's recommended, generic, performance-based disability assessment tool. It is structured around six basic functioning domains:

- D1: Cognition – understanding & communicating
- D2: Mobility – moving & getting around
- D3: Self-care – hygiene, dressing, eating & staying alone
- D4: Getting along – interacting with other people
- D5: Life activities – domestic responsibilities, leisure, work & school
- D6: Participation – joining in community activities

The "clinical" version of the WHODAS questionnaire collects information about functioning and problems in functioning – i.e., disability – by means of a face-to-face interview conducted by a trained interviewer who asks a set of standardized questions and, if necessary, follow-up probe questions. WHODAS uses a 5-level response scale (1 = None, 2 = Mild, 3 = Moderate, 4 = Severe, 5 = Extreme or Cannot do) to rate each question. In extraordinary circumstances (e.g., COVID lockdown), WHODAS can be administered in a telephone or video interview by the trained professional. Respondents are informed that their answers about each domain of functioning should adopt the perspective of performance, i.e., that they should describe what they do taking into account the experiences in their daily life and the environmental barriers and facilitators they experience. For the pilot, the 36-item version of WHODAS was chosen to create a full picture

of the disability experienced by the respondent in their everyday life. The 36 WHODAS items are summarised in Table 1 by functioning domain.

Table 1: WHODAS items for the 36-item long form

	In the past 30 days, how much difficulty did you have in:
	Understanding and communicating
D1.1	Concentrating on doing something for ten minutes?
D1.2	Remembering to do important things?
D1.3	Analysing and finding solutions to problems in day-to-day life?
D1.4	Learning a new task, for example, learning how to get to a new place?
D1.5	Generally understanding what people say?
D1.6	Starting and maintaining a conversation?
	Getting around
D2.1	Standing for long periods such as 30 minutes?
D2.2	Standing up from sitting down?
D2.3	Moving around inside your home?
D2.4	Getting out of your home?
D2.5	Walking a long distance such as a kilometre [or equivalent]?
	Self-care
D3.1	Washing your whole body?
D3.2	Getting dressed?
D3.3	Eating?
D3.4	Staying by yourself for a few days?
	Getting along with people
D4.1	Dealing with people you do not know?
D4.2	Maintaining a friendship?
D4.3	Getting along with people who are close to you?
D4.4	Making new friends?
D4.5	Sexual activities?
	Life activities
D5.1	Taking care of your household responsibilities?
D5.2	Doing most important household tasks well?
D5.3	Getting all the household work done that you needed to do?
D5.4	Getting your household work done as quickly as needed?
D5.5	Your day-to-day work/school?
D5.6	Doing your most important work/school tasks well?
D5.7	Getting all the work done that you need to do?
D5.8	Getting your work done as quickly as needed?
	Participation in society:
D6.1	How much of a problem did you have in joining in community activities as anyone else can?
D6.2	How much of a problem did you have because of barriers or hindrances in the world around you?
D6.3	How much of a problem did you have living with dignity cause of attitudes and actions of others?
D6.4	How much time did you spend on your health condition or its consequences?
D6.5	How much have you been emotionally affected by your health condition?
D6.6	How much has your health been a drain on the financial resources of you or your family?
D6.7	How much of a problem did your family have because of your health problems?
D6.8	How much of a problem did you have in doing things by yourself for relaxation or pleasure?

2.1. Sample characteristics

A total of 3307 individuals participated in the WHODAS survey. The data for 65 individuals were not included in the analyses for this report, given a very high number of missing values in their responses (> 20%). The socio-demographic characteristics of the remaining N = 3242 individuals are shown in Table 2. Participants were all between 18 and 68 years old, i.e., of working age, and capable of understanding and responding to the questions asked by the interviewer. The proportion of male participants was lower (45.1% vs. female share of 54.9%). The average age was 50.6 years (SD (standard deviation) = 11.9). Most participants indicated their marital status as being married (51.6%), 26.3% indicated never having been married. Most respondents were living independently in the community (92.3%). Many participants indicated either having paid work (39.7%) or being unemployed for either health (21.6%) or other reasons (15.6%).

Table 2: Pilot sample – descriptive statistics for the full sample

Distribution of the full sample across selected socio-demographic characteristics (in per cent)

N	3242
Gender = Male (%)	1460 (45.1)
Age – mean (SD)	50.63 (11.91)
Years of Education – mean (SD)	11.34 (3.65)
Marital Status (%)	
Never married	854 (26.3)
Currently married	1673 (51.6)
Separated	205 (6.3)
Divorced	239 (7.4)
Widowed	114 (3.5)
Cohabiting	156 (4.8)
Living Condition (%)	
Independent in the community	2979 (92.3)
Assisted living	241 (7.5)
Hospitalized	6 (0.2)
Work Status (%)	
Paid work	1287 (39.7)
Self-employed	186 (5.7)
Non-paid work	9 (0.3)
Student	107 (3.3)
Keeping house	241 (7.4)
Retired	177 (5.5)
Unemployed (health reasons)	699 (21.6)
Unemployed (other reasons)	506 (15.6)
Other	28 (0.9)

The sample's socio-demographic information by region is presented in Table 3. For this pilot evaluation, data was available for four Italian regions, i.e., Campania, Sardinia, Lombardy, and Trentino, with most data from Campania and Lombardy. The proportion of male participants was below 50% for all four regions with no significant differences across the four regions. Instead, mean ages differed significantly across regions, with 52.2 years (SD = 10.9) in Campania, 49.8 years (SD = 12.3) in Lombardy, 50.7 years (SD = 13.1) in Sardinia, and 48.8 years (SD = 12.5) in Trentino. An average of about 11 years of education was reported for all regions. Across the regions, most participants indicated their marital status as being married (61% in Campania, 47.5% in Lombardy, 41.8% in Sardinia, 43.5% in Trentino). Most respondents were living independently in the community. The percentage of individuals living in assisted living was highest in Trentino (15.7%). The data on employment was collected in different manners so that for some of the data collected in Campania detailed information is missing, i.e., it was not possible to determine if unemployment was health-related or not or if the work activity was for an employer or self-employed. The proportion of the participants in paid work was especially high in Lombardy (47.7%) and Trentino (48.2%).

Table 3: Pilot sample – descriptive statistics for each of the four participating regions

Distribution of the four regional samples across selected socio-demographic characteristics (in per cent)

	Campania	Lombardy	Sardinia	Trentino
N	1223	1327	182	510
Gender = Male (%)	543 (44.5)	580 (43.7)	86 (47.3)	251 (49.2)
Age – mean (SD)	52.24 (10.89)	49.81 (12.25)	50.71 (13.08)	48.84 (12.48)
Years Education – mean (SD)	11.38 (3.81)	11.26 (3.59)	11.32 (3.99)	11.46 (3.29)
Marital Status (%)				
Never married	260 (21.3)	356 (26.8)	66 (36.3)	172 (33.7)
Currently married	745 (61.0)	630 (47.5)	76 (41.8)	222 (43.5)
Separated	81 (6.6)	79 (6.0)	14 (7.7)	31 (6.1)
Divorced	64 (5.2)	124 (9.3)	12 (6.6)	39 (7.6)
Widowed	46 (3.8)	48 (3.6)	6 (3.3)	14 (2.7)
Cohabiting	26 (2.1)	90 (6.8)	8 (4.4)	32 (6.3)
Living Condition (%)				
Independent in community	1166 (96.6)	1206 (90.9)	182 (100.0)	425 (83.3)
Assisted living	41 (3.4)	120 (9.0)	0 (0.0)	80 (15.7)
Hospitalized	0 (0.0)	1 (0.1)	0 (0.0)	5 (1.0)
Work Status (%)				
Paid work	358 (29.3)	633 (47.7)	50 (27.5)	246 (48.2)
Self-employed	94 (7.7)	65 (4.9)	7 (3.8)	20 (3.9)
Non-paid work	2 (0.2)	4 (0.3)	0 (0.0)	3 (0.6)
Student	25 (2.0)	57 (4.3)	11 (6.0)	14 (2.7)
Keeping house	139 (11.4)	65 (4.9)	18 (9.9)	19 (3.7)
Retired	64 (5.2)	76 (5.7)	14 (7.7)	23 (4.5)
Unemployed (health reasons)	209 (17.1)	299 (22.5)	60 (33.0)	131 (25.7)
Unemployed (other reasons)	324 (26.5)	123 (9.3)	21 (11.5)	38 (7.5)
Other	7 (0.6)	4 (0.3)	1 (0.5)	16 (3.1)

Table 4 presents the frequency and percentages of observed ICD-11 diagnostic chapters, with the caveat that the data on health conditions were collected differently in the four regions. The different coding systems included the ICD-9 as well as the Italian DM 1992 condition coding system. Health condition codes were semi-manually linked to the closest ICD-11 chapter; the latest version of the WHO's International Classification of Diseases (ICD-11). Many people in the data set have more than one diagnosis. If several diagnoses would link to just one ICD chapter, the chapter was reported only once. For example, if an individual had three codes that all described aspects of his or her circulatory problem, it was only mentioned once that he or she was having a "Disease of the circulatory system (code 11)". The situation is different for people with more than one condition from different ICD chapters. Information regarding the importance or priority of different diagnoses, i.e., whether a condition is the persons main health problem or a secondary comorbidity, was unavailable for most data. It was decided to include in the following analyses by health condition all ICD-chapter diagnoses recorded for a person. Such, the total number of conditions presented in Table 4 is larger than the total sample as a person with two different conditions would be counted twice. This should not affect the findings by health condition.

The four most and roughly equally prevalent health conditions were: diseases of the musculoskeletal system and connective tissue (N = 578, 16.5% of the sample), diseases of the circulatory system (N = 564, 16.1%), neoplasms (N = 558; 15.9%); and mental and behavioural disorders (N = 535; 15.3%). In total, these four conditions accounted for almost two-thirds of the entire sample.

Disaggregated by region, however, the prevalence of reported conditions varied considerably (Table 5). Neoplasms had the highest prevalence in Campania (22.1%). Diseases of the musculoskeletal system and diseases of connective tissue were most prevalent in Sardinia (26.9%), with the other three regions reporting less than 20% of conditions classified in this chapter. Mental, behavioural, or neurodevelopmental disorders were reported by 19% in Campania, 17.9% in Trentino, 11.6% in Lombardy, and 10.7% in Sardinia. Sardinia also reported 11.1% of endocrine, nutritional, or metabolic diseases. The somewhat different diagnostic distribution in Sardinia may in part be explained by the relatively small regional sample.

Table 4: Prevalence of diagnoses in the study population by ICD-11 health condition chapter

ICD-Chapter	N	%
1 Certain infectious or parasitic diseases	14	0.4 %
2 Neoplasms	558	15.93 %
3 Diseases of the blood or blood-forming organs	6	0.17 %
4 Diseases of the immune system	36	1.03 %
5 Endocrine, nutritional or metabolic diseases	155	4.42 %
6 Mental, behavioural or neurodevelopmental disorders	535	15.27 %
8 Diseases of the nervous system	281	8.02 %
9 Diseases of the visual system	87	2.48 %
10 Diseases of the ear or mastoid process	115	3.28 %
11 Diseases of the circulatory system	564	16.1 %
12 Diseases of the respiratory system	150	4.28 %
13 Diseases of the digestive system	138	3.94 %
14 Diseases of the skin	2	0.06 %
15 Diseases of the musculoskeletal system and diseases of connective tissue	578	16.5 %
16 Diseases of the genitourinary system	50	1.43 %
20 Development anomalies	14	0.4 %
21 Symptoms, signs or clinical findings, not elsewhere classified	51	1.46 %
22 Injury, poisoning or certain other consequences of external causes	22	0.63 %
24 Factors influencing health status or contact with health services	147	4.2 %
Total	3503	100%

Note: The total exceeds the sample size as more than one ICD-11 chapter can be reported per individual.

Table 5 Prevalence of diagnoses by ICD-11 chapter for the four participating regions

	Campania		Lombardy		Sardinia		Trentino	
ICD-Chapter	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
1 Certain infectious or parasitic diseases	1	0.09%	5	0.36%	3	1.39%	5	0.6%
2 Neoplasms	234	22.12%	229	16.4%	14	6.48%	81	9.72%
3 Diseases of the blood or blood-forming organs	2	0.19%	2	0.14%	2	0.93%	0	0%
4 Diseases of the immune system	4	0.38%	30	2.15%	0	0%	2	0.24%
5 Endocrine, nutritional or metabolic diseases	59	5.58%	44	3.15%	24	11.11%	28	3.36%
6 Mental, behavioural or neurodevelopmental disorders	201	19%	162	11.6%	23	10.65%	149	17.89%
8 Diseases of the nervous system	38	3.59%	139	9.96%	18	8.33%	86	10.32%
9 Diseases of the visual system	16	1.51%	34	2.44%	5	2.31%	32	3.84%
10 Diseases of the ear or mastoid process	25	2.36%	57	4.08%	3	1.39%	30	3.6%
11 Diseases of the circulatory system	188	17.77%	184	13.18%	31	14.35%	161	19.33%
12 Diseases of the respiratory system	19	1.8%	94	6.73%	2	0.93%	35	4.2%
13 Diseases of the digestive system	26	2.46%	68	4.87%	12	5.56%	32	3.84%
14 Diseases of the skin	0	0%	0	0%	2	0.93%	0	0%
15 Diseases of the musculoskeletal system and diseases of connective tissue	191	18.05%	217	15.54%	58	26.85%	112	13.45%
16 Diseases of the genitourinary system	12	1.13%	21	1.5%	7	3.24%	10	1.2%
20 Development anomalies	3	0.28%	6	0.43%	2	0.93%	3	0.36%
21 Symptoms, signs or clinical findings, not elsewhere classified	10	0.95%	23	1.65%	6	2.78%	12	1.44%
22 Injury, poisoning or certain other consequences of external causes	4	0.38%	13	0.93%	1	0.46%	4	0.48%
24 Factors influencing health status or contact with health services	25	2.36%	68	4.87%	3	1.39%	51	6.12%
	1058	100%	1396	100%	216	100%	833	100%

2.1. WHODAS frequencies

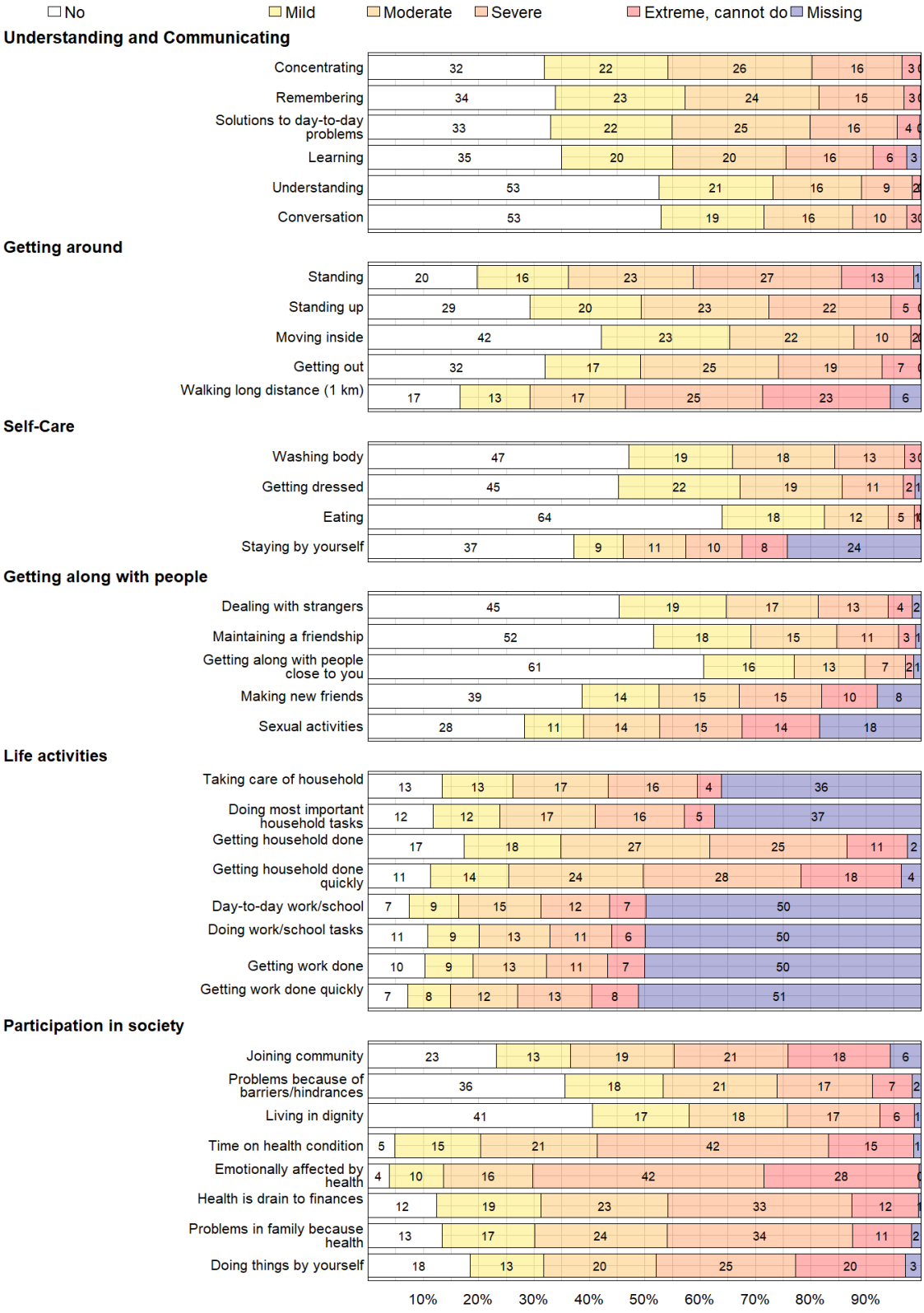
Table 6 shows the descriptive statistics for the 36 WHODAS items, including the number and percentage of missing values. More than half of the participants indicated their problems to be severe or extreme for the following items: D6.5 - How much have you been emotionally affected by your health condition? (69.8%), D6.4 - How much time did you spend on your health condition or its consequences? (57.2%), D2.5 - Walking a long distance, such as a kilometre [or equivalent]? (47.9%), D5.4 - Getting your household work done as quickly as needed? (46.6%), D6.6 - How much has your health been a drain on the financial resources of you or your family? (45.2%), D6.8 - How much of a problem did you have in doing things by yourself for relaxation or pleasure? (44. 9%), and D6.7 – How much of a problem did your family have because of your health problems? (44.2%). The descriptive statistics for the ratings of the WHODAS disaggregated by the four participating regions of Italy are reported in the Annex.

Table 6: Frequencies and percentages of WHODAS responses for the full sample

Item	No	Mild	Moderate	Severe	Extreme, cannot do	Missing
D1.1	1033 (31.86%)	723 (22.3%)	845 (26.06%)	528 (16.29%)	107 (3.3%)	6 (0.19%)
D1.2	1098 (33.87%)	759 (23.41%)	785 (24.21%)	497 (15.33%)	100 (3.08%)	3 (0.09%)
D1.3	1069 (32.97%)	711 (21.93%)	810 (24.98%)	510 (15.73%)	134 (4.13%)	8 (0.25%)
D1.4	1132 (34.92%)	654 (20.17%)	663 (20.45%)	511 (15.76%)	195 (6.01%)	87 (2.68%)
D1.5	1706 (52.62%)	665 (20.51%)	518 (15.98%)	298 (9.19%)	53 (1.63%)	2 (0.06%)
D1.6	1715 (52.9%)	605 (18.66%)	518 (15.98%)	319 (9.84%)	83 (2.56%)	2 (0.06%)
D2.1	638 (19.68%)	535 (16.5%)	734 (22.64%)	865 (26.68%)	423 (13.05%)	47 (1.45%)
D2.2	948 (29.24%)	652 (20.11%)	746 (23.01%)	717 (22.12%)	175 (5.4%)	4 (0.12%)
D2.3	1368 (42.2%)	750 (23.13%)	727 (22.42%)	334 (10.3%)	61 (1.88%)	2 (0.06%)
D2.4	1036 (31.96%)	561 (17.3%)	807 (24.89%)	609 (18.78%)	225 (6.94%)	4 (0.12%)
D2.5	538 (16.59%)	412 (12.71%)	557 (17.18%)	803 (24.77%)	751 (23.16%)	181 (5.58%)
D3.1	1529 (47.16%)	604 (18.63%)	599 (18.48%)	411 (12.68%)	98 (3.02%)	1 (0.03%)
D3.2	1469 (45.31%)	708 (21.84%)	601 (18.54%)	359 (11.07%)	67 (2.07%)	38 (1.17%)
D3.3	2076 (64.03%)	596 (18.38%)	375 (11.57%)	151 (4.66%)	37 (1.14%)	7 (0.22%)
D3.4	1205 (37.17%)	289 (8.91%)	369 (11.38%)	327 (10.09%)	266 (8.2%)	786 (24.24%)
D4.1	1470 (45.34%)	629 (19.4%)	538 (16.59%)	412 (12.71%)	138 (4.26%)	55 (1.7%)
D4.2	1674 (51.63%)	569 (17.55%)	502 (15.48%)	361 (11.14%)	102 (3.15%)	34 (1.05%)
D4.3	1964 (60.58%)	531 (16.38%)	416 (12.83%)	237 (7.31%)	49 (1.51%)	45 (1.39%)
D4.4	1255 (38.71%)	450 (13.88%)	471 (14.53%)	481 (14.84%)	327 (10.09%)	258 (7.96%)
D4.5	917 (28.29%)	343 (10.58%)	449 (13.85%)	480 (14.81%)	457 (14.1%)	596 (18.38%)
D5.1	435 (13.42%)	413 (12.74%)	560 (17.27%)	520 (16.04%)	142 (4.38%)	1172 (36.15%)
D5.2	382 (11.78%)	391 (12.06%)	556 (17.15%)	526 (16.22%)	177 (5.46%)	1210 (37.32%)
D5.3	562 (17.33%)	568 (17.52%)	870 (26.84%)	807 (24.89%)	354 (10.92%)	81 (2.5%)
D5.4	365 (11.26%)	459 (14.16%)	789 (24.34%)	924 (28.5%)	588 (18.14%)	117 (3.61%)
D5.5	241 (7.43%)	290 (8.95%)	482 (14.87%)	400 (12.34%)	213 (6.57%)	1616 (49.85%)
D5.6	349 (10.76%)	302 (9.32%)	413 (12.74%)	363 (11.2%)	198 (6.11%)	1617 (49.88%)
D5.7	334 (10.3%)	279 (8.61%)	430 (13.26%)	359 (11.07%)	218 (6.72%)	1622 (50.03%)
D5.8	232 (7.16%)	250 (7.71%)	396 (12.21%)	432 (13.33%)	274 (8.45%)	1658 (51.14%)
D6.1	753 (23.23%)	431 (13.29%)	610 (18.82%)	668 (20.6%)	597 (18.41%)	183 (5.64%)
D6.2	1153 (35.56%)	576 (17.77%)	665 (20.51%)	562 (17.33%)	231 (7.13%)	55 (1.7%)
D6.3	1316 (40.59%)	567 (17.49%)	573 (17.67%)	543 (16.75%)	201 (6.2%)	42 (1.3%)
D6.4	158 (4.87%)	502 (15.48%)	682 (21.04%)	1357 (41.86%)	498 (15.36%)	45 (1.39%)
D6.5	123 (3.79%)	318 (9.81%)	523 (16.13%)	1357 (41.86%)	908 (28.01%)	13 (0.4%)
D6.6	400 (12.34%)	613 (18.91%)	745 (22.98%)	1077 (33.22%)	388 (11.97%)	19 (0.59%)
D6.7	435 (13.42%)	540 (16.66%)	776 (23.94%)	1086 (33.5%)	348 (10.73%)	57 (1.76%)
D6.8	597 (18.41%)	433 (13.36%)	659 (20.33%)	817 (25.2%)	640 (19.74%)	96 (2.96%)

Figure 1 visualizes how the items of the WHODAS questionnaire have been rated. The percentage of missing values was highest, i.e., about 50%, for items D5.5 to D5.8 that assess difficulties at work (or in school), as all participants were over 18 years old, with many being unemployed for health or other reasons. More than 30% of missing values were also found for D5.1 - Taking care of household responsibilities and D5.2 - Doing most important household tasks, as these two questions were not consistently assessed across all the regions at the start of the assessment pilot.

Figure 1: Percentage of ratings by degree of civil disability for each WHODAS item



2.2. WHODAS score distribution

Figure 2 shows the distribution of the total raw scores obtained when adding up the 32 items of WHODAS. The total WHODAS score does not include items D5.5 to D5.8 because of their large number of missing values. The total raw WHODAS score ranges from 32 to 160, although a few total scores below 32 are possible as the scores are computed on the raw data with some missing values (less than 20%). Coloured segments in Figure 2 indicate the position and value of the 1^{st}, 2^{nd}, and 3^{rd} quartiles, with a median score (2^{nd} quartile) of 75. The density lines in Figure 3 show the density of the observed scores (red line) and the corresponding normal distribution with the same mean and standard deviation (dotted line). Scores in this sample for Italy are distributed relatively normally, which was a common finding also in other countries where WHODAS was pilot tested (Latvia, Lithuania, Greece, Bulgaria, Romania, Seychelles).

Figure 2: Raw score distribution of the WHODAS

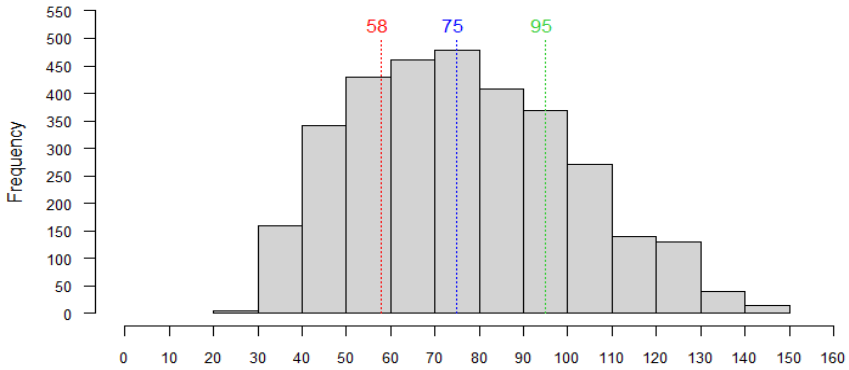
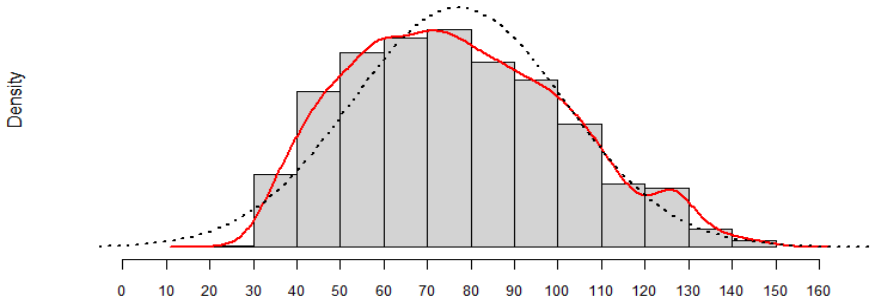
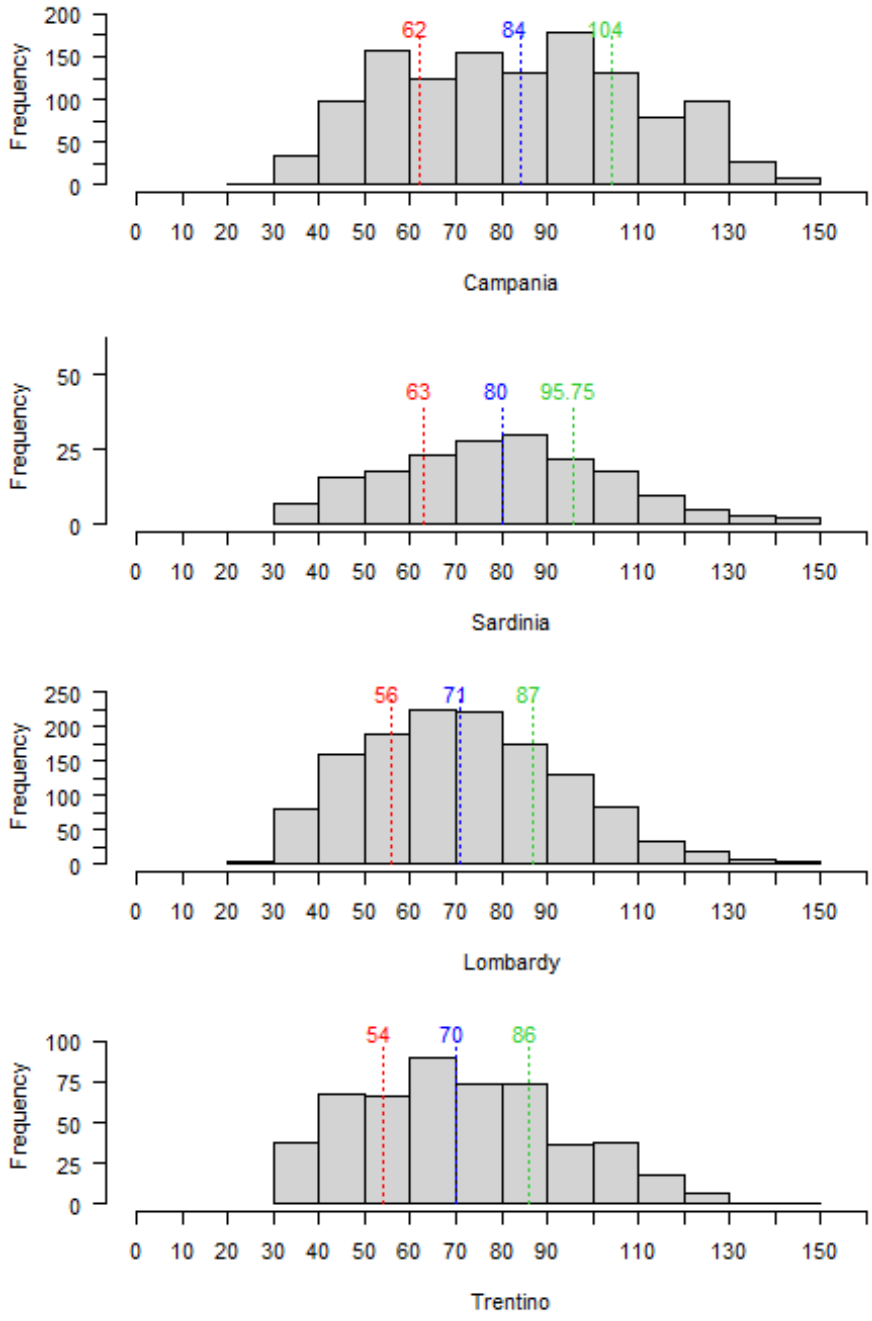


Figure 3: Score density: observed density and random normal density



The distributions of the WHODAS raw scores in the four regions that participated in the pilot present small differences (Figure 4). The highest median WHODAS score (blue raw dotted line) is found for Campania (Q2 = 84) and the lowest median score in Trentino (Q2 = 70). Higher WHODAS raw scores indicate higher levels of disability. Otherwise, however, the figures show rather normally distributed WHODAS raw scores for all four participating regions.

Figure 4: Raw score distributions of the WHODAS in the four regions



3. Psychometric analysis

Rasch analysis is a statistical method from the field of probabilistic measurement. It is a modern test theory approach first introduced in the 1960's by the Danish mathematician George Rasch (Rasch, 1960). The items of the WHODAS are rated by means of more than two response options and were so calibrated with the Partial Credit Model (Masters, 1982), an extension of the Rasch model developed for dichotomous data.

Rasch analysis is essentially testing several measurement assumptions (Bond and Fox, 2001; Tennant and Conaghan, 2007): (1) the targeting of the scale, (2) the model reliability, (3) the ordering of the items' response options, (4) the absence of correlation between items (Local Item Dependencies - LID), (5) the fit of the items to the Rasch model, (6) the absence of effects of person factors such as gender and age on item responses (Differential Item Functioning - DIF), and (7) the unidimensionality of the questionnaire. If these measurement assumptions can be met, a questionnaire can be considered psychometrically sound and derived total scores therefore be considered interval-scaled and operative for measurement.

For a well-performing questionnaire, it is expected that the difficulty of the items is matched to the level of ability of the measured population, i.e., the questionnaire should not be too easy or too difficult. Statistically, good targeting (*assumption #1*) is achieved if the mean item difficulty and mean person ability are approximating zero. A Person Separation Index (PSI) above 0.8 speaks for a good reliability of the scale and values above 0.9 for very good reliability (*assumption #2*). The PSI indicates how well the scale can discriminate levels of functioning in the population. The Cronbach α, which is typically also reported, is a classical measure of the internal consistency of the data, i.e., how well the items work to describe one construct (Nunnally and Bernstein, 1994). In the presence of disordered response options (*assumption #3*), an analysis of response probability curves allows to determine which response options cause problems and decide on strategies to aggregate disordered response options. For example, if for an item the response options 2 and 1 appear reversed and indicate that an increase of difficulty cannot be discriminated, the item responses can be recoded so that these options represent only one level of response. LID often occurs when items are redundant and measure approximately the same aspect of a construct (*assumption #4*). The most widely reported statistic for the item dependencies is the Q3 matrix, i.e., the correlation matrix of the Rasch residuals (Yen, 1984). Residual correlations above 0.2 are considered as not acceptable and a way to address these local item dependencies, without deleting items, is to aggregate (i.e., to sum up) the correlating items into so-called testlets (Yen, 1993). In item testlets, the ordering of the thresholds is not expected anymore. For good item fit (*assumption #5*), the infit and outfit values are expected below 1.2 (Smith, Schumacker, and Bush, 1998). The outfit statistic is more sensitive to outliers as the infit statistic. Ideally, items of a questionnaire should be fair and not favour sample subgroups. The analysis of DIF allows to flag exogenous variables, or DIF variables (*assumption #6*), which conduct to a lack of invariance of the item difficulty (Holland and Wainer, 1993). It is worthwhile to note that a DIF analysis is not always indicating a metric bias but can also simply represent subgroups with unequal underlying ability (Boone, Staver and Yale, 2014). DIF analysis was conducted for age and gender, to determine the items which are sensitive to those external covariates. Finally, a questionnaire should measure only one construct. If a questionnaire shows to have several separate dimensions, the validity of one summary total score is not supported. Unidimensionality (*assumption #7*) was assessed with a principal component analysis of the Rasch residuals (Smith, 2002). Typically, a first eigenvalue lower than 1.8 is deemed indicative of unidimensionality. Based on simulation analyses, Smith and Miao (1994) suggested considering the size of the second component instead, with values below 1.4 indicative of unidimensionality. The above analyses were all performed with the software R (Team, 2016).

3.1. Metric properties of WHODAS

The pilot implemented the complete 36-item version of WHODAS. However, items D5.5 to D5.8, which were answered only by persons working or in education, presented about 50% of missing values and were not used in the psychometric analysis. The Items D5.1 - Taking care of your household responsibilities? (36.2%) and D5.2 - Doing most important household tasks well? (37.3%) also showed high proportions of missing values as they were not collected in the first assessment wave in Lombardy. Somewhat higher proportions of missing values were also found for D3.4 - Staying by yourself for a few days (24.2%) and D4.5 – Sexual activities (18.4%). The remaining items contained <10% of missing values, see Table 6 above. Items D5.5 to D5.8, on education and employment, were removed from the analysis and the WHODAS-based functioning score was built with the remaining 32 items. With a sample size of N = 3242, it is assumed that the Rasch model can be run reliably (Fellinghauer, Prodinger, and Tennant, 2018); nonetheless, the person parameter scores were computed using an imputed dataset (Stekhoven and Buhlmann, 2012).

The Rasch analysis showed that the scale is multidimensional, with a strong tendency of the items to load (i.e., to correlate with other variables) within WHODAS domains. Only a few items loaded across domains and, similarly, only a few items were free of dependencies. To solve the issues of multidimensionality and local-item dependencies, correlating items were aggregated by accounting for the domain structure of the WHODAS questionnaire. The detailed resulting statistics are shown in Table 7 for the reliability and quality of targeting, in Table 8 for the fit statistics at the start of the analysis, and in Table 9 for the fit statistics after having made necessary adjustments. Findings can be summarised as follows:

(1) The population included in this analysis presented a very good targeting to the scale (Table 7).

(2) The item reliability was high but also inflated at the beginning of the analysis because of item dependencies ($PSI = 0.95$, Cronbach $\alpha = 0.95$). Reliability was still found to be good also after the adjustments were made ($PSI = 0.88$, Cronbach $\alpha = 0.89$).

(3) The response thresholds of 23 of the 32 items of the WHODAS questionnaire presented disordering (Figure 5). Locally dependent items can be an explanation for the disordering, as well as a lack of discrimination between the two first response options, i.e., answer categories “None” and “Mild”. Figure 7 shows the range of assessment with the testlets that aggregate the items by domains.

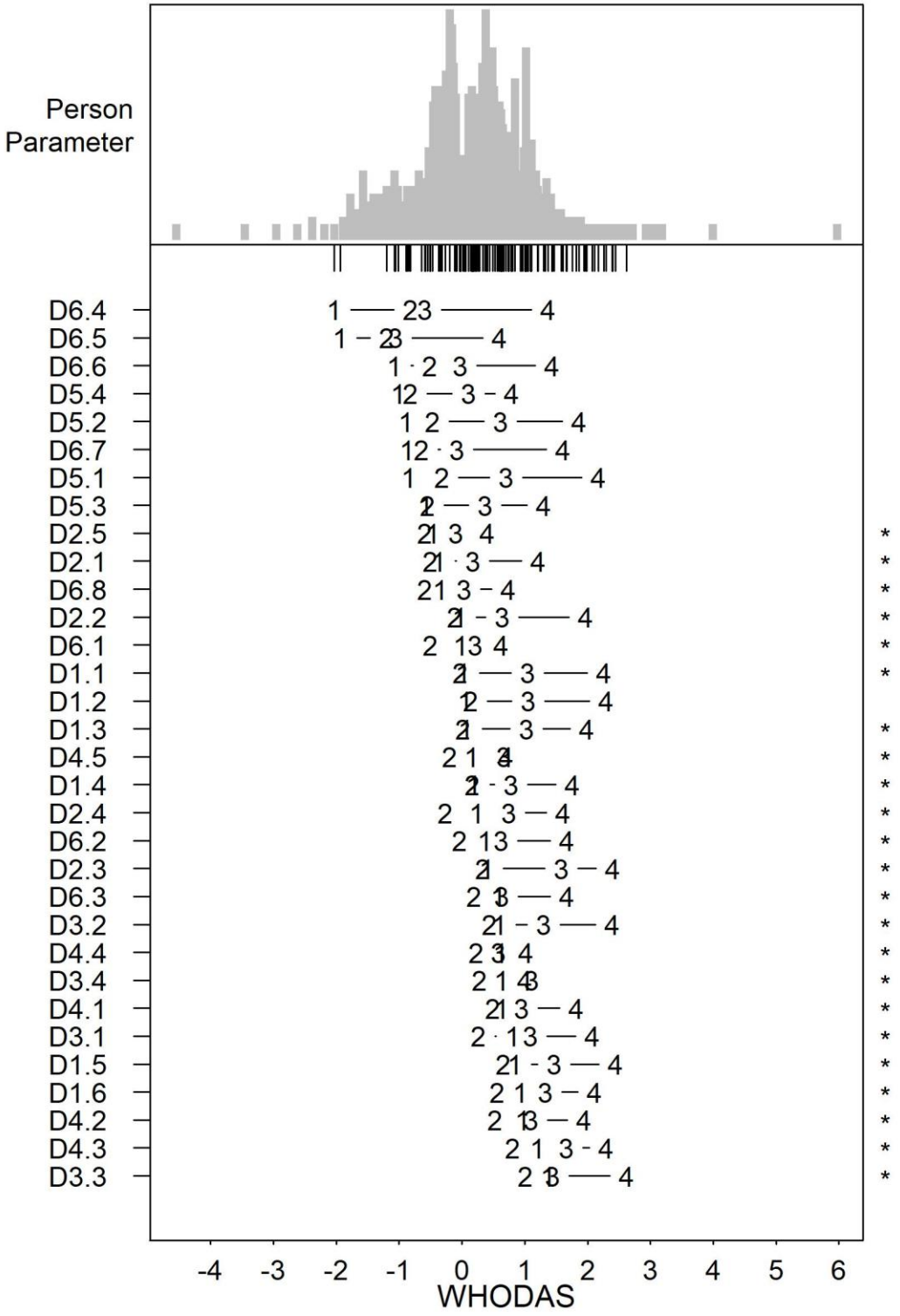
(4) The analysis of the residual dependencies showed strong local dependencies among most items of the WHODAS questionnaire (Figure 6), with a tendency of questionnaire items from the same domain to associate. To address these dependencies, items were aggregated considering the domain structure of the tool. The thresholds of the testlets are not expected to be ordered.

(5) The item fit is good if the infit and outfit values are below 1.2. Three out of the 32 items showed misfit with infit or outfit above the cut-off: D1.5 - Generally understanding what people say?, D6.4 - How much time did you spend on your health condition or its consequences, and D6.6 - How much has your health been a drain on the financial resources of you or your family?. After aggregation of the items by domain, all testlets showed good infit and outfit values, below 1.2.

(6) The DIF analysis indicated mainly that the WHODAS-domains are sensitive to age groups. Further, responses to domain 1 (Understanding and communicating) and domain 5a (Life activities) are affected by the gender of the respondent.

(7) The principal component analysis indicated that the items cluster by domains which results in multidimensionality, with a very high 1^{st} eigenvalue of 5.29 and a 2^{nd} eigenvalue of 2.87. After adjustments, i.e., aggregation of items by WHODAS domains, the 1^{st} eigenvalue dropped to 1.93 and the 2^{nd} eigenvalue to 1.29, indicating unidimensionality according to the defined criteria.

Figure 5: Person item map before collapsing of response options



*indicate disordered thresholds

Table 7: Targeting and Reliability of WHODAS items

	Targeting			
	Start Mean	SD	Final Mean	SD
Difficulty	0.56	0.95	0.21	0.55
Ability	0.00	0.88	0.00	0.40
	PSI	**Alpha**	**PSI**	**Alpha**
Reliability	0.95	0.95	0.88	0.89

Figure 6: Local Item Dependencies before creation of testlets

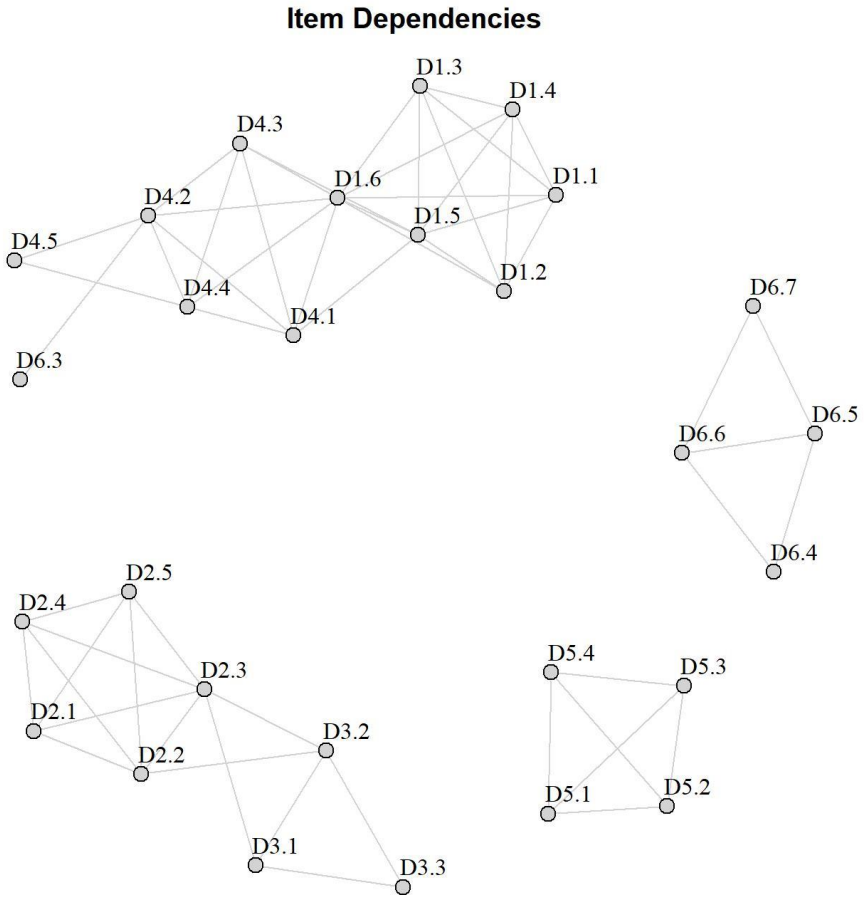


Figure 7: Person item map after solving dependencies by aggregating items by domains

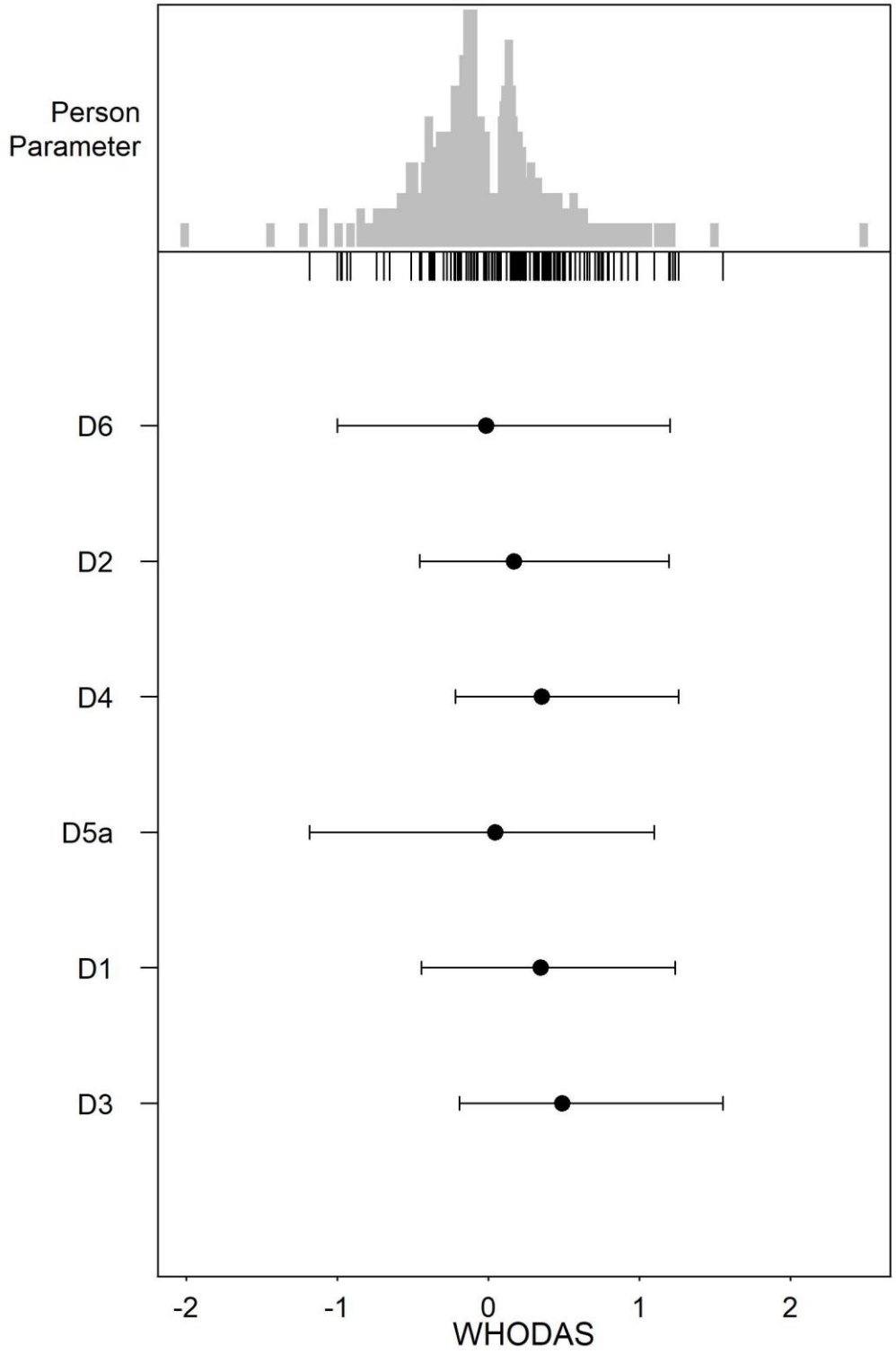


Table 8: WHODAS properties at start: item difficulties, fit (outfit and infit), local item dependencies, and differential item functioning

WHODAS Item Nbr.	Outfit[1]	Infit[1]	Item Difficulty	Disordered Thresholds	LID[3]
D1.1	1.01	1.01	0.82	x	D1.2, D1.3, D1.4, D1.5, D1.6
D1.2	1.04	1.03	0.88		D1.1, D1.3, D1.4, D1.5, D1.6
D1.3	0.91	0.92	0.77	x	D1.1, D1.2, D1.4, D1.5, D1.6
D1.4	1.01	0.98	0.72	x	D1.1, D1.2, D1.3, D1.5, D1.6
D1.5	*1.3*	1.12	1.35	x	D1.1, D1.2, D1.3, D1.4, D1.6, D4.1, D4.3
D1.6	1.13	1.07	1.23	x	D1.1, D1.2, D1.3, D1.4, D1.5, D4.1, D4.2, D4.3, D4.4
D2.1	1.03	1.02	0.13	x	D1.5, D1.6, D4.2, D4.3, D4.4
D2.2	1.1	1.07	0.61	x	D1.5, D1.6, D4.1, D4.2, D4.4
D2.3	0.8	0.87	1.18	x	D1.6, D4.1, D4.3, D4.4, D4.5, D6.3
D2.4	0.73	0.76	0.58	x	D1.6, D4.1, D4.2, D4.3, D4.5
D2.5	1.03	1.02	-0.19	x	D2.2, D2.3, D2.4, D2.5
D3.1	0.77	0.86	1.06	x	D2.1, D2.3, D2.4, D2.5, D3.2
D3.2	0.76	0.85	1.18	x	D2.1, D2.2, D2.4, D2.5, D3.1, D3.2
D3.3	1.12	1.09	1.61	x	D2.1, D2.2, D2.3, D2.5
D3.4	0.89	0.9	0.75	x	D2.1, D2.2, D2.3, D2.4
D4.1	0.98	1.01	0.97	x	D2.2, D2.3, D3.1, D3.3
D4.2	0.91	0.97	1.13	x	D2.3, D3.2, D3.3
D4.3	1	1.02	1.49	x	D3.1, D3.2
D4.4	0.94	0.97	0.61	x	
D4.5	1.11	1.1	0.33	x	D4.2, D4.4
D5.1	0.65	0.66	0.43		D4.2
D5.2	0.65	0.65	0.28		D5.2, D5.3, D5.4
D5.3	0.69	0.71	0.14		D5.1, D5.3, D5.4
D5.4	0.73	0.74	-0.23		D5.1, D5.2, D5.4
D6.1	0.88	0.91	0.08	x	D5.1, D5.2, D5.3
D6.2	0.99	1.02	0.66	x	
D6.3	1.17	1.15	0.77	x	
D6.4	*1.28*	*1.25*	-0.52		D6.5, D6.6
D6.5	1.07	1.08	-0.90		D6.4, D6.6, D6.7
D6.6	*1.32*	*1.29*	-0.04		D6.4, D6.5, D6.7
D6.7	1.07	1.07	0.00		D6.5, D6.6
D6.8	0.98	0.99	-0.04	x	

1 Infit and Outfit expected below 1.2 for the absence of underfit
2 In testlets, i.e., aggregated locally dependent items, the ordering of thresholds is not expected anymore
3 Local item dependency (LID) significant with r > (mean(Q3) + 0.2)

Table 9: WHODAS properties after the necessary adjustment: item difficulties, fit (outfit and infit), local item dependencies, and differential item functioning

WHODAS Item Nbr.	Label	Outfit1	Infit1	Item Difficulty	Disordered Thresholds	LID3
Testlet 1	D1.1-D1.6	1.07	1.09	0.35	n.a.2	no
Testlet 2	D2.1-D2.5	0.89	0.91	0.17	n.a.2	no
Testlet 3	D3.1-D3.4	0.65	0.7	0.49		no
Testlet 4	D4.1-D4.5	0.85	0.91	0.35	n.a.2	no
Testlet 5	D5.1-D5.4	0.69	0.7	0.04		no
Testlet 6	D6.1-D6.8	0.75	0.75	-0.01	n.a.2	no

1 Infit and Outfit expected below 1.2 for the absence of underfit
2 In testlets, i.e. aggregated locally dependent items, the ordering of thresholds is not expected anymore
3 Local item dependency (LID) significant with r > (mean(Q3) + 0.2)).

Statistical psychometric testing confirmed the validity and reliability of the WHODAS instrument in the Italian context and environment. Statistical analysis of the psychometric properties of WHODAS with the data piloted in Italy shows that functioning data collected with WHODAS display robust psychometric properties. It is important to keep in mind that the WHO developed WHODAS explicitly to statistically capture the construct of functioning from the perspective of performance – i.e., the experience of performing activities by a person with an underlying health problem in their everyday life environment. There is an abundance of evidence from the scientific literature – further supported by the results of this pilot – that WHODAS is a psychometrically sound instrument that reliably and validly collects information about levels of disability.

Based on satisfactory psychometric properties, one can confidently conclude that information collected with the WHODAS questionnaire is robust, viable, and relevant and that it validly represents the construct of disability as understood in the ICF and the UN Convention on the Rights of Persons with Disabilities (UNCRPD). Including the WHODAS questionnaire into disability status assessment in Italy would therefore (i) significantly strengthen the method of assessment currently in use (which is a medical assessment based on the existence of impairments) and align it with Italy's general approach to disability; (ii) bring it closer to the ICF and UNCRPD understanding of disability; and (iii) harmonise the approach to assessment with the ICF functioning-based approach used in subsequent individual needs assessments.

The psychometric analysis of the pilot data from Campania, Lombardy, Sardinia, and Trentino demonstrates the validity and reliability of the WHODAS instrument as a tool to measure disability also in Italy. More than this, all four regions have been able to implement the pilot successfully and, importantly, the pilot has shown the ability of social workers to take the lead in carrying out the necessary clinical interviews ("clinical" in this case meaning that the WHODAS questionnaire was implemented by clinically trained social workers).

3.2. WHODAS score transformation

WHODAS raw scores are transformed into a scale of 0 (no disability) to 100 points (complete disability). Table 10 shows the score transformation derived from the data collected in Italy. The table includes both logit-scale Rasch ability estimates and rescaled "user-friendly" scores on a scale from 0 to 100. The table allows recoding scores from the 32 WHODAS items into a psychometrically sound interval-scaled metric.

Table 10: WHODAS score transformation table

WHODAS Score	Rasch Logit	0-100 Score	WHODAS* Score	Rasch* Logit	0-100* Score
32	-1.95	0	96	0.18	48
33	-1.44	11	97	0.19	48
34	-1.24	16	98	0.2	48
35	-1.11	19	99	0.22	48
36	-1.01	21	100	0.23	49
37	-0.94	23	101	0.24	49
38	-0.87	24	102	0.25	49
39	-0.81	25	103	0.26	49
40	-0.76	27	104	0.28	50
41	-0.72	28	105	0.29	50
42	-0.67	28	106	0.3	50
43	-0.64	29	107	0.32	50
44	-0.6	30	108	0.33	51
45	-0.57	31	109	0.34	51
46	-0.54	31	110	0.35	51
47	-0.51	32	111	0.37	52
48	-0.48	33	112	0.38	52
49	-0.46	33	113	0.4	52
50	-0.44	34	114	0.41	53
51	-0.41	34	115	0.43	53
52	-0.39	35	116	0.44	53
53	-0.37	35	117	0.46	54
54	-0.35	36	118	0.47	54
55	-0.33	36	119	0.49	54
56	-0.32	36	120	0.5	55
57	-0.3	37	121	0.52	55
58	-0.28	37	122	0.54	55
59	-0.27	38	123	0.55	56
60	-0.25	38	124	0.57	56
61	-0.24	38	125	0.59	57
62	-0.22	39	126	0.61	57
63	-0.21	39	127	0.63	57
64	-0.19	39	128	0.65	58
65	-0.18	39	129	0.67	58
66	-0.17	40	130	0.69	59
67	-0.15	40	131	0.71	59
68	-0.14	40	132	0.73	60
69	-0.13	41	133	0.75	60
70	-0.11	41	134	0.77	61
71	-0.1	41	135	0.79	61
72	-0.09	41	136	0.82	62
73	-0.08	42	137	0.84	62
74	-0.07	42	138	0.86	63
75	-0.06	42	139	0.88	63
76	-0.04	42	140	0.91	64
77	-0.03	43	141	0.93	64
78	-0.02	43	142	0.96	65
79	-0.01	43	143	0.98	65
80	0	44	144	1.01	66
81	0.01	44	145	1.04	67
82	0.02	44	146	1.06	67
83	0.04	44	147	1.09	68
84	0.05	44	148	1.12	69
85	0.06	45	149	1.16	69
86	0.07	45	150	1.19	70
87	0.08	45	151	1.23	71
88	0.09	45	152	1.27	72
89	0.1	46	153	1.32	73
90	0.11	46	154	1.35	73
91	0.12	46	155	1.38	74
92	0.14	46	156	1.45	76
93	0.15	47	157	1.6	79
94	0.16	47	158	1.85	85
95	0.17	47	159	2.17	92
96	0.18	48	160	2.54	100

4. Comparing WHODAS scores and civil invalidity ratings

4.1. Meaningful cut-off points

There are, to our knowledge, no agreed and published cut-offs for the WHODAS score that would be applicable to a population with diverse health conditions to categorize the severity of their disability. Having established cut-offs would allow to easily detect individuals with significant disabilities and to reflect and, eventually, reconsider attributed civil invalidity percentages. Some studies report the 90th or 95th percentile of the WHODAS score distribution as being the best cut-off to diagnose severe disability or dysfunctionality in some specific groups, such as post-partum women (Mayrink et al., 2018) or the elderly population (Ferrer et al., 2019). A minimal clinically important difference in scores for the WHODAS has not been established yet (Federici et al., 2017). However, based on several previous and comparable pilot projects conducted by the World Bank using the WHODAS questionnaire, in Greece, Latvia, Lithuania, and Bulgaria, meaningful WHODAS disability cut-off points for the Rasch-based 0-100 score are suggested as follows:

- Score 0-25: No functioning restrictions (i.e., no difficulties in performance/disability)
- Score 26-40: Moderate functioning restrictions (i.e., moderate difficulties in performance/disability)
- Score 41-60: Severe functioning restrictions (i.e., severe difficulties in performance/disability)
- Score 61-100: Very severe functioning restrictions (i.e., very severe difficulties in performance/disability)

A score of 40 would thus be a central cut-off for determining the presence of a disability and, thus, eligibility for services. In total, the sample presented N = 74 (2.3%) of individuals having no functioning restrictions, N = 972 (30.0%) of individuals with moderate functioning restrictions, N = 2120 (65.4%) of individuals with severe functioning restrictions, and N = 76 (2.3%) of individuals with very severe functioning restrictions.

Later in the report, additional cut-offs are introduced to split the two middle groups in which most people are concentrated – thereby distinguishing lower and higher moderate functioning restrictions (with WHODAS scores of 26-34 and 35-40, respectively) as well as lower and higher severe functioning restrictions (with WHODAS scores of 41-48 and 49-60, respectively).

The civil invalidity percentages attributed to persons with health problems in Italy, following the assessment, can be divided into different categories in various ways. While there are no cut-off points for a discretionary assessment, entitlement for a number of benefits and supports suggest the following as a meaningful split:

- 0-33%: no invalidity
- 34-66%: moderate invalidity, of which
  - 34-45%: lower moderate invalidity
  - 46-66%: higher moderate invalidity
- 67-99%: severe invalidity, of which
  - 67-73%: lower severe invalidity
  - 74-99%: higher severe invalidity

- 100%: very severe invalidity

In total, the pilot sample presented N = 81 (2.8%) of individuals with no civil invalidity, N = 1129 (38.8%) of individuals with moderate civil invalidity, N = 1076 (37%) with severe civil invalidity, and N = 623 (21.4%) of individuals with very severe civil invalidity rated as 100%. There were N = 333 (10.3%) individuals in the data set with no reported civil invalidity percentage. Further dividing the middle groups, according to the above-suggested scale, returns N = 225 (7.7%) of individuals with lower moderate invalidity and N = 904 (31.1%) with higher moderate invalidity. Likewise, N = 420 individuals (14.4%) with lower severe invalidity and N = 656 (22.6%) of individuals with higher severe invalidity. The different levels of invalidity are key to obtaining supports from Italy's social protection system. For example, with a civil invalidity percentage above 46% individuals can request employment support, with more than 67% prostheses are provided free of charge, and with more than 74% people can receive a non-contributory disability allowance.

4.2. Sample characteristics according to cut-off points

Table 11 presents the socio-demographic characteristics of the sample disaggregated by level of disability based on the WHODAS score. With 68.9%, the percentage of men was higher in the group with no disability and close to or below 50% otherwise. There is a statistically significant increase in mean age (p-value < 0.001) across disability levels from 45.7 years with no disability to 53.5 years with very severe disability. The average number of years of education decreases significantly with increasing disability status (p-value < 0.001) from about 12 years with no disability to about 11 years with very severe disability. With regard to the living situation, 77.3% of participants with very severe disability lived independently in the community, with shares above 90% for all other groups. The percentage of persons in paid work decreased from 56.8% in the group with no disability to 21.1% for those with very severe disability.

Table 11: Sample descriptive statistics by disability severity based on the WHODAS questionnaire

	No	Moderate	Severe	Very severe
N	74	972	2120	76
Gender = Male (%)	51 (68.9)	491 (50.5)	884 (41.8)	34 (44.7)
Age – mean (SD)	45.74 (15.98)	49.32 (12.35)	51.29 (11.52)	53.45 (9.53)
Years of Education – mean (SD)	12.05 (3.54)	11.75 (3.67)	11.14 (3.64)	10.96 (3.42)
Living Condition (%)				
Independent in the community	73 (98.6)	936 (96.6)	1912 (90.7)	58 (77.3)
Assisted living	1 (1.4)	33 (3.4)	190 (9.0)	17 (22.7)
Hospitalized	0 (0.0)	0 (0.0)	6 (0.3)	0 (0.0)
Marital Status (%)				
Never married	31 (41.9)	273 (28.1)	531 (25.0)	19 (25.0)
Currently married	33 (44.6)	506 (52.1)	1094 (51.6)	40 (52.6)
Separated	4 (5.4)	54 (5.6)	142 (6.7)	5 (6.6)
Divorced	3 (4.1)	58 (6.0)	170 (8.0)	8 (10.5)
Widowed	0 (0.0)	27 (2.8)	86 (4.1)	1 (1.3)
Cohabiting	3 (4.1)	53 (5.5)	97 (4.6)	3 (3.9)
Work Status (%)				
Paid work	42 (56.8)	474 (48.8)	755 (35.6)	16 (21.1)
Self-employed	7 (9.5)	71 (7.3)	107 (5.1)	1 (1.3)
Non-paid work	0 (0.0)	3 (0.3)	6 (0.3)	0 (0.0)
Student	5 (6.8)	46 (4.7)	56 (2.6)	0 (0.0)
Keeping house	1 (1.4)	68 (7.0)	167 (7.9)	5 (6.6)
Retired	8 (10.8)	38 (3.9)	119 (5.6)	12 (15.8)
Unemployed (health reasons)	4 (5.4)	122 (12.6)	540 (25.5)	33 (43.4)
Unemployed (other reasons)	7 (9.5)	142 (14.6)	349 (16.5)	8 (10.5)
Other	0 (0.0)	8 (0.8)	19 (0.9)	1 (1.3)

Table 12 presents the socio-demographic characteristics of the sample disaggregated by the level of civil invalidity, following the above-proposed cut-off categories. The percentage of men is higher and above 50% only in the group of persons with no civil invalidity. Again, there is a statistically significant increase in the mean age (p-value < 0.001) across degrees of civil invalidity, from 45.2 years in the group with no invalidity to 52.9 years in the group with very severe civil invalidity. The average number of years of education is slightly above 11 years across all invalidity levels. The share of people living independently in the community is about 85.2% among those with very severe invalidity and above 90% for the other groups. Finally, the percentage of persons in paid work decreases from about 44.4% in the group of persons with no or moderate civil invalidity to 32.4% in the group of persons with very severe disability.

Table 12: Sample descriptive statistics by impairment severity based on assessment of civil invalidity

	No	Minor	Moderate	Severe
N	81	1129	1076	623
Gender = Male (%)	44 (54.3)	498 (44.1)	507 (47.2)	271 (43.6)
Age – mean (SD)	45.16 (14.20)	48.94 (12.38)	51.69 (11.59)	52.87 (10.72)
Years of Education – mean (SD)	11.41 (3.22)	11.37 (3.61)	11.27 (3.78)	11.52 (3.64)
Living Condition (%)				
Independent in the community	76 (93.8)	1074 (95.8)	996 (93.1)	529 (85.2)
Assisted living	5 (6.2)	45 (4.0)	73 (6.8)	89 (14.3)
Hospitalized	0 (0.0)	2 (0.2)	1 (0.1)	3 (0.5)
Marital Status (%)				
Never married	28 (34.6)	301 (26.7)	277 (25.7)	156 (25.0)
Currently married	35 (43.2)	583 (51.7)	557 (51.8)	326 (52.3)
Separated	2 (2.5)	79 (7.0)	73 (6.8)	39 (6.3)
Divorced	9 (11.1)	79 (7.0)	72 (6.7)	47 (7.5)
Widowed	1 (1.2)	32 (2.8)	47 (4.4)	24 (3.9)
Cohabiting	6 (7.4)	54 (4.8)	50 (4.6)	31 (5.0)
Work Status (%)				
Paid work	36 (44.4)	509 (45.1)	397 (36.9)	202 (32.4)
Self-employed	6 (7.4)	69 (6.1)	58 (5.4)	37 (5.9)
Non-paid work	1 (1.2)	5 (0.4)	3 (0.3)	0 (0.0)
Student	7 (8.6)	50 (4.4)	24 (2.2)	15 (2.4)
Keeping house	5 (6.2)	87 (7.7)	78 (7.3)	44 (7.1)
Retired	2 (2.5)	23 (2.0)	67 (6.2)	67 (10.8)
Unemployed (health reasons)	14 (17.3)	204 (18.1)	238 (22.1)	163 (26.2)
Unemployed (other reasons)	10 (12.3)	171 (15.1)	200 (18.6)	88 (14.1)
Other	0 (0.0)	11 (1.0)	10 (0.9)	7 (1.1)

4.3. Pathologies, WHODAS scores and civil invalidity ratings

Table 13 presents the mean WHODAS score, on the 0-100 scale, disaggregated by health conditions, and as a memorandum item also the distribution of the population across the ICD-11 chapters. The 24 individuals having a health condition linked to “21 Symptoms, signs or clinical findings that could not be classified elsewhere” presented the highest mean WHODAS score of 46.66 (SD = 11.4). The least disabling conditions as measured by the WHODAS score are development anomalies with a mean score of 40.8 (SD = 8); these were reported by 14 pilot participants. Among the four most frequent pathologies, “mental, behavioural or neurodevelopmental disorders” has the highest mean WHODAS score (45, SD = 8), while the other three (neoplasms, circulatory system, musculoskeletal system) all have mean scores at or close to 43.

Table 13: Frequency and percentage of ICD chapters and corresponding mean and standard deviation (SD) of the WHODAS scores

	N	Mean (SD)
1 Certain infectious or parasitic diseases	14 (0.4%)	43.4 (9.75)
2 Neoplasms	558 (15.93%)	43.44 (8.08)
3 Diseases of the blood or blood-forming organs	6 (0.17%)	48.47 (8.07)
4 Diseases of the immune system	36 (1.03%)	45.13 (7.77)
5 Endocrine, nutritional or metabolic diseases	155 (4.42%)	43.85 (7.67)
6 Mental, behavioural or neurodevelopmental disorders	535 (15.27%)	44.95 (7.99)
8 Diseases of the nervous system	281 (8.02%)	45.09 (8.34)
9 Diseases of the visual system	87 (2.48%)	41.49 (8.56)
10 Diseases of the ear or mastoid process	115 (3.28%)	42.02 (8.07)
11 Diseases of the circulatory system	564 (16.1%)	42.65 (7.94)
12 Diseases of the respiratory system	150 (4.28%)	41.38 (9.07)
13 Diseases of the digestive system	138 (3.94%)	43.16 (7.22)
14 Diseases of the skin	2 (0.06%)	43.12 (10.7)
15 Diseases of the musculoskeletal system and diseases of connective tissue	578 (16.5%)	43.51 (7.13)
16 Diseases of the genitourinary system	50 (1.43%)	41.64 (8.66)
20 Development anomalies	14 (0.4%)	40.8 (8)
21 Symptoms, signs or clinical findings, not elsewhere classified	51 (1.46%)	46.66 (11.39)
22 Injury, poisoning, or other consequences of external causes	22 (0.63%)	42.87 (5.99)

Table 14 disaggregates the sample by pathology and degree of civil invalidity. By and large, the results show that mean WHODAS scores tend to increase with the invalidity degree for most pathologies although the results must be interpreted with caution, due to the small number of cases in the group with no invalidity (N = 81). It is not the same condition that consistently receives the highest WHODAS rating across the different civil invalidity degree groups. Looking at the four main pathologies only, for which the sample size is large enough to draw reliable conclusions, the following can be observed:

- Diseases of the musculoskeletal system are the dominant pathology among people with a moderate level of civil invalidity (25.5% of those with degrees 34-66%). For those diseases, mean WHODAS scores clearly and gradually increase with the invalidity degree, from around 38.1 to 49.8.
- Neoplasms are the dominant pathology among people with very severe levels of invalidity (38.5% of those with a degree of 100%). Mean WHODAS scores are lower than for the other main diseases, at all invalidity levels with degrees above 33%.
- Diseases of the circulatory system are particularly frequent in the two middle invalidity categories, moderate and severe disability (i.e., degree 34-99%). Mean WHODAS scores generally lie between those for neoplasms and for diseases of the musculoskeletal system.
- The percentage of mental, behavioural, or neurodevelopmental disorders increases slightly with an increasing invalidity degree, with a high WHODAS mean compared to the other main diseases.
- The mean WHODAS scores increase with the invalidity degree for all four main pathologies.

Table 14: Frequency and percentage of ICD chapters by civil invalidity degree and mean and SD of the corresponding WHODAS score

	No invalidity (0-33%)		Moderate invalidity (34-66%)		Severe invalidity (67-99%)		Very severe invalidity (100%)	
	N no	mean (SD) no	N minor	Mean (SD) minor	N moderate	Mean (SD) moderate	N severe	Mean (SD) severe
1 Certain infectious or parasitic diseases	1 (1.25%)	43.67 (-)	5 (0.4%)	39.87 (13.52)	6 (0.43%)	43.87 (6.99)	2 (0.27%)	50.65 (9.53)
2 Neoplasms	5 (6.25%)	42.09 (18.48)	74 (5.86%)	38.7 (7.75)	190 (13.58%)	41.61 (7.35)	285 (38.46%)	45.91 (7.56)
3 Diseases of the blood or blood-forming organs			1 (0.08%)	44.41 (-)	2 (0.14%)	42.64 (1.8)	2 (0.27%)	50.53 (9.7)
4 Diseases of the immune system	1 (1.25%)	51.6 (-)	14 (1.11%)	41.74 (7.58)	16 (1.14%)	44.69 (4.91)	5 (0.67%)	54.73 (9.18)
5 Endocrine, nutritional or metabolic diseases	2 (2.5%)	32.31 (11.66)	44 (3.48%)	40.49 (8.85)	77 (5.5%)	44.33 (6.67)	30 (4.05%)	48.38 (5.19)
6 Mental, behavioural or neurodevelopmental disorders	8 (10%)	37.72 (11.26)	169 (13.38%)	42.79 (7.06)	258 (18.44%)	44.92 (7.98)	99 (13.36%)	49.34 (7.37)
8 Diseases of the nervous system	7 (8.75%)	39.03 (9.9)	87 (6.89%)	40.59 (7.48)	101 (7.22%)	45.03 (7.52)	85 (11.47%)	50.27 (7)
9 Diseases of the visual system	5 (6.25%)	32.28 (14.01)	35 (2.77%)	40.73 (8.42)	37(2.64%)	43.69 (7.66)	10 (1.35%)	40.62 (6.31)
10 Diseases of the ear or mastoid process	9 (11.25%)	40.8 (7.17)	64 (5.07%)	41.1 (7.94)	34 (2.43%)	43.97 (8.14)	8 (1.08%)	42.48 (9.6)
11 Diseases of the circulatory system	6 (7.5%)	42.14 (6.63)	220 (17.42%)	40.81 (7.42)	269 (19.23%)	42.98 (7.76)	65 (8.77%)	47.04 (8.48)
12 Diseases of the respiratory system	2 (2.5%)	39.82 (1.03)	86 (6.81%)	39.61 (10.05)	53 (3.79%)	43.31 (7.19)	9 (1.21%)	47.24 (4.91)
13 Diseases of the digestive system	2 (2.5%)	49.76 (2.59)	43 (3.4%)	40.73 (6.43)	68 (4.86%)	42.32 (6.88)	25 (3.37%)	49.1 (6.32)
14 Diseases of the skin					1 (0.07%)	50.69 (-)	1 (0.13%)	35.55 (-)
15 Diseases of the musculoskeletal system	20 (25%)	38.05 (5.72)	322 (25.49%)	42.3 (7.07)	183 (13.08%)	44.57 (6.45)	47 (6.34%)	49.75 (6.16)
16 Diseases of the genitourinary system	1 (1.25%)	35.97 (-)	14 (1.11%)	39.7 (5.58)	19 (1.36%)	38.79 (7.86)	16 (2.16%)	47.08 (9.82)
20 Development anomalies			5 (0.4%)	37.69 (6.71)	6 (0.43%)	39.95 (6.4)	3 (0.4%)	47.66 (11.25)
21 Symptoms not elsewhere classified	3 (3.75%)	27.98 (11.49)	9 (0.71%)	38.74 (7.04)	19 (1.36%)	45.95 (7.46)	20 (2.7%)	53.69 (10.79)
22 Injury, poisoning or other external causes	2 (2.5%)	42.87 (2.88)	9 (0.71%)	44.01 (6.94)	7 (0.5%)	38.72 (4.65)	3 (0.4%)	47.1 (1.88)

Table 15 looks at the mean WHODAS score and the mean civil invalidity percentage per ICD chapter, comparing the situation when the linked health condition chapter appeared as standalone diagnostical information versus when it was reported in addition to other health condition chapters; thereby comparing cases of single morbidity with cases of comorbidity.

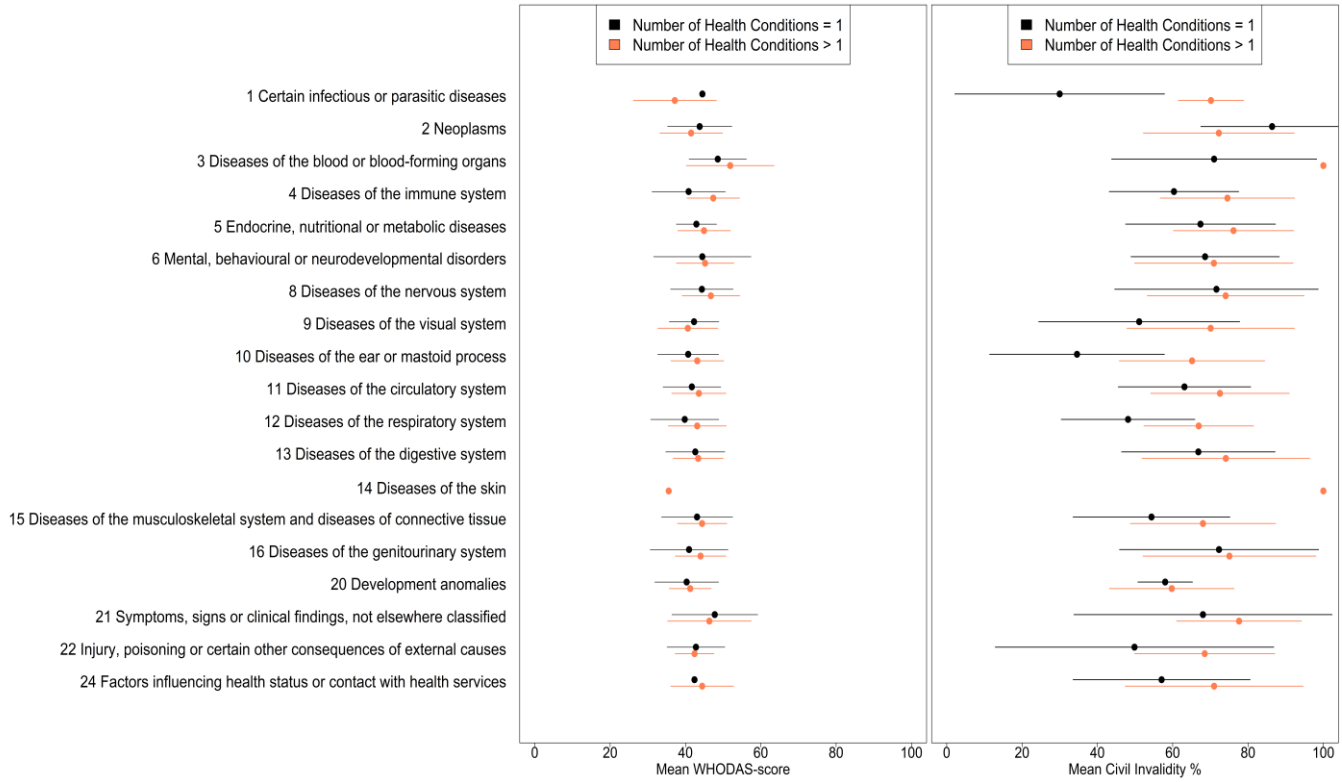
The average WHODAS score per ICD chapter hardly changes whether it is a single diagnosis or part of multiple diagnoses. In contrast, the average civil invalidity percentage is in many cases higher when a person is diagnosed with multiple conditions. In other words, the WHODAS score per ICD chapter varies significantly less than the civil invalidity percentage: it appears that co-morbidity has an influence on the civil invalidity percentage but not on the WHODAS score. The data do not allow an interpretation of this finding but the discretionary freedom in the civil invalidity assessment could play a role, i.e., assessors perceiving people with co-morbidity as more disabled – a finding not corroborated by WHODAS scores.

Table 15: Mean and standard deviation of the WHODAS score and the civil invalidity percentage per ICD chapter: comparing results for single diagnoses with cases of comorbidity

	Number ICD chapter linked = 1 (i.e., single diagnosis)			Number ICD chapter linked > 1 (i.e., multiple diagnoses)		
	N	WHODAS score mean (SD)	Civil invalidity percentage mean (SD)	N	WHODAS score mean (SD)	Civil invalidity percentage mean (SD)
1 Certain infectious or parasitic diseases	3	44.49 (0.76)	30 (27.84)	6	37.16 (11.01)	70.17 (8.68)
2 Neoplasms	430	43.78 (8.09)	86.38 (18.93)	58	41.51 (8.31)	72.26 (20.05)
3 Diseases of the blood and its organs	3	48.57 (7.64)	71 (27.22)	2	51.88 (11.62)	100 (NA)
4 Diseases of the immune system	12	40.85 (9.04)	60.33 (17.17)	15	47.37 (6.97)	74.53 (17.83)
5 Endocrine or nutritional diseases	67	42.91 (7.85)	67.38 (19.89)	52	44.93 (6.97)	76.15 (15.95)
6 Mental or behavioural disorders	316	44.47 (8.09)	68.61 (19.68)	97	45.22 (7.61)	70.94 (21.06)
8 Diseases of the nervous system	117	44.36 (9.43)	71.62 (27.01)	71	46.76 (7.67)	74.06 (20.83)
9 Diseases of the visual system	30	42.28 (10.32)	51.1 (26.68)	48	40.63 (7.94)	70.08 (22.25)
10 Diseases of the ear or mastoid process	30	40.73 (8.54)	34.6 (23.2)	65	43.13 (7)	65.14 (19.31)
11 Diseases of the circulatory system	236	41.71 (8.45)	63.13 (17.57)	186	43.56 (7.17)	72.55 (18.4)
12 Diseases of the respiratory system	55	39.81 (11.37)	48.15 (17.74)	57	43.11 (7.69)	66.91 (14.56)
13 Diseases of the digestive system	57	42.61 (7.67)	66.81 (20.39)	46	43.39 (6.62)	74.11 (22.28)
15 Diseases of the musculoskeletal system	299	43.08 (7.61)	54.37 (20.82)	177	44.42 (6.56)	68.04 (19.26)
16 Diseases of the genitourinary system	23	40.96 (9.74)	72.3 (26.43)	19	44.02 (6.73)	75.05 (22.96)
20 Development anomalies	4	40.33 (5.34)	58 (7.26)	4	41.26 (5.55)	59.75 (16.56)
21 Symptoms not elsewhere classified	21	47.78 (12.92)	68.05 (34.27)	19	46.35 (11.1)	77.63 (16.56)
22 Injury, poisoning, other external causes	7	42.79 (8.29)	49.86 (36.97)	13	42.41 (5.14)	68.5 (18.6)
24 Factors influencing health status or contact with health services	58	42.4 (6.58)	57.05 (23.5)	51	44.44 (8.31)	71 (23.61)

Figure 8 shows the same result graphically by comparing the mean WHODAS score and the mean civil invalidity percentage as well as their standard deviation, per ICD-chapter, for people with one and people with more than one condition. The black dots and segments represent the WHODAS score or civil invalidity percentage for ICD chapters where individuals had only a single diagnosis. The red dots represent the ICD chapters where individuals had multiple diagnoses in addition to the diagnosis linked to that chapter. The segments to the left and right of the mean represent standard deviations (not confidence intervals). The results visually repeat the finding presented in Table 15: while WHODAS scores do not seem to depend on the number of diagnoses, civil invalidity percentages systematically increase for cases of comorbidity.

Figure 8: Mean and standard deviation of the WHODAS score and the civil invalidity percentage per ICD chapter: graphical comparison between single diagnoses and cases of comorbidity



The following figures pursue the comparison between the disability score based on the WHODAS questionnaire and the result of the civil invalidity assessment. Figure 9 looks at the distribution of WHODAS scores against the distribution of civil invalidity percentages. While WHODAS disability scores are distributed normally around a mean of 43.2, with a standard deviation of 8.45, civil invalidity percentages seem to be distributed erratically, with higher frequencies at distinct locations on the continuum linked with critical cut-offs for eligibility for specific social benefits and services. The discretionary method of assigning invalidity percentages with limited guidelines and standards might explain the concentration at the cut-offs. In practice, this turns the invalidity scale into an ordinal scale with just a few possible outcomes.

Figure 9: Distribution of the WHODAS score and of the civil invalidity percentage

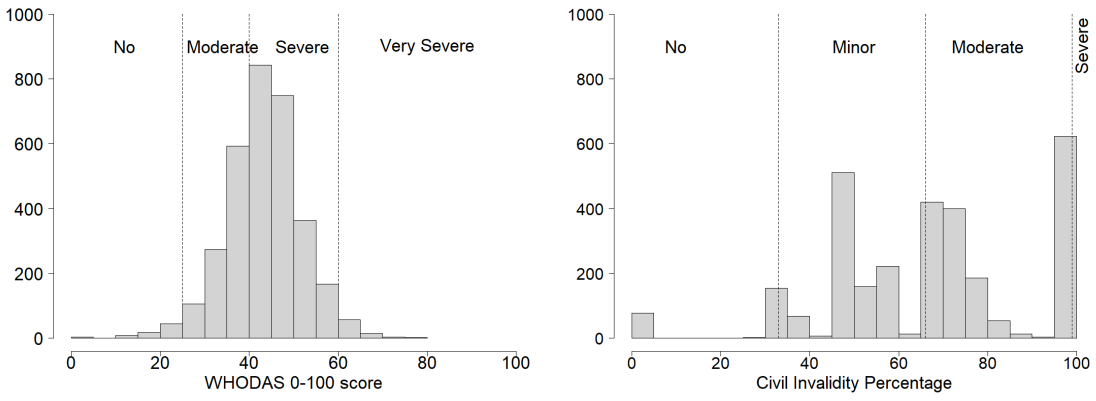


Table 16 shows the four civil invalidity groups disaggregated by WHODAS disability groups. In interpreting these findings, it is important to keep in mind that a moderate civil invalidity level should not be understood to be equal to a moderate disability level. These are two different perspectives, correlated only moderately: WHODAS measures lived experience of disability in the person's everyday environment; civil invalidity assesses disability based on health condition/impairment (medical approach). The table shows that the number of individuals that fall in opposite severity groups is negligible: there is only one person with a very severe WHODAS disability but no civil invalidity and no one with very severe civil invalidity and no WHODAS disability. However, less extreme seemingly contradictory cases are more frequent: there are, for example, 94 persons with very severe civil invalidity and only moderate WHODAS disability. Likewise, the data include 40 persons with severe WHODAS disability but no civil invalidity.

Table 16: Frequencies of civil invalidity degree groups by WHODAS disability group

		Civil Invalidity Degree Groups				
		No	Moderate	Severe	Very Severe	Missing
WHODAS Disability Groups	No	10 (0.31%)	39 (1.2%)	14 (0.43%)	0 (0%)	11 (0.34%)
	Moderate	30 (0.93%)	434 (13.39%)	298 (9.19%)	94 (2.9%)	116 (3.58%)
	Severe	40 (1.23%)	646 (19.93%)	754 (23.26%)	481 (14.84%)	199 (6.14%)
	Very severe	1 (0.03%)	10 (0.31%)	10 (0.31%)	48 (1.48%)	7 (0.22%)
	Missing	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Table 17 presents the civil invalidity degree groups, with a finer division of the moderate and severe groups of civil invalidity, in line with the system-specific cut-off points in Italy. The distribution echoes the above findings but also shows how important these system-specific cut-offs are in the assessment decisions. Among those who exhibit both moderate WHODAS disability and moderate civil invalidity, three in four persons fall under the "higher moderate" group. Similarly, among those who exhibit both severe WHODAS disability and severe civil invalidity, two in three persons fall under the "higher severe" group.

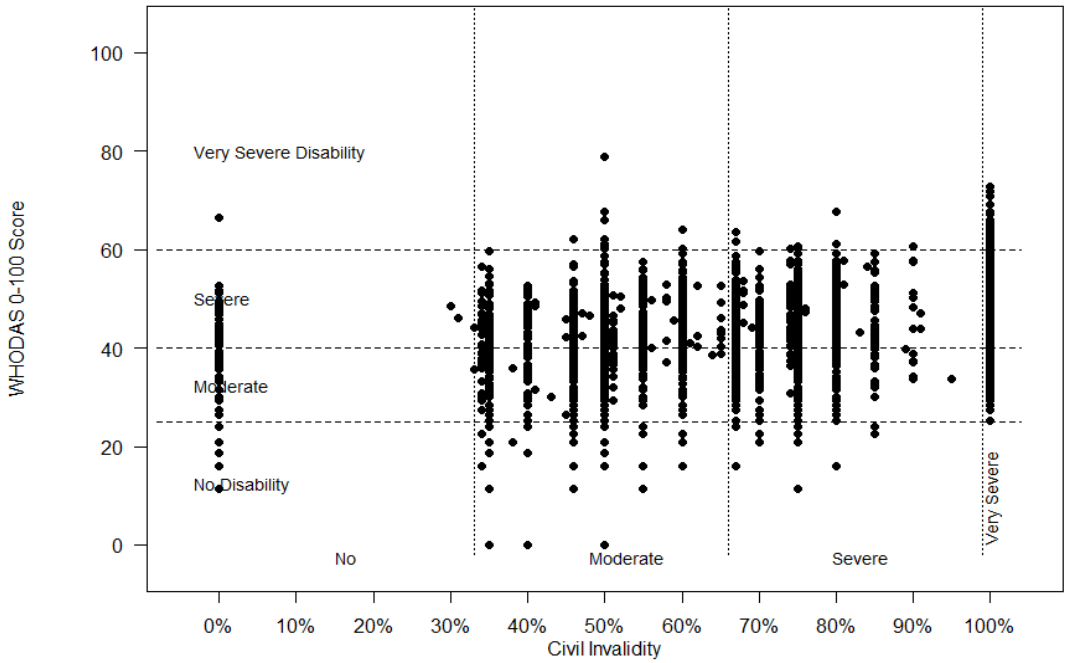
Table 17: Frequencies of civil invalidity degree groups with additional sub-categories by WHODAS disability group

		Civil Invalidity Degree Groups						
		No	Lower Moderate	Higher Moderate	Lower Severe	Higher Severe	Very Severe	Missing
WHODAS Disability Groups	No	10 (0.31%)	14 (0.43%)	25 (0.77%)	4 (0.12%)	10 (0.31%)	0 (0%)	11 (0.34%)
	Moderate	30 (0.93%)	103 (3.18%)	331 (10.21%)	132 (4.07%)	166 (5.12%)	94 (2.9%)	116 (3.58%)
	Severe	40 (1.23%)	108 (3.33%)	538 (16.6%)	281 (8.7%)	473 (14.6%)	481 (14.8%)	199 (6.1%)
	Very severe	1 (0.03%)	0 (0%)	10 (0.31%)	3 (0.09%)	7 (0.22%)	48 (1.48%)	7 (0.22%)
	Missing	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Figure 10 also compares the distribution of individual civil invalidity percentages and WHODAS scores. The figure shows the full distribution of data points for the WHODAS score (y-axis) and the civil invalidity percentage (x-axis). Horizontal lines represent the cut-offs for the WHODAS score, from no disability to moderate, severe, and very severe disability, and vertical lines represent the cut-offs for the civil invalidity percentage (again, no, moderate, severe, and very severe). The two scores show a positive correlation but only at a very moderate level (R = 0.33). This is expected because disability cannot be inferred from medical conditions/impairment only: two individuals with the same medical diagnosis will be assigned the same percentage of disability based on medical criteria for the assessment. However, they may experience very different levels of disability (functioning limitation and participation restrictions or performance in the ICF disability understanding) depending on their environment.

Some notable exceptions can be observed on the plot, such as individuals having 0% of civil invalidity while reporting moderate to very severe disability according to the WHODAS questionnaire looking at their functioning levels across different life domains. Similarly, some individuals with a civil invalidity percentage above 66% (i.e., with severe or very severe invalidity) are found not to have any disability based on their WHODAS score.

Figure 10: WHODAS score distributions at respective civil-invalidity cut-offs: full sample



Figures 10a-10d show that the results are similar for all regions although Sardinia stands out as a region in which both civil invalidity ratings and WHODAS scores are more concentrated in the middle (e.g., most individuals have minor to moderate civil invalidity and none of them presented very severe disability based on WHODAS). The resulting correlation between WHODAS scores and civil invalidity ratings is lowest in Trentino (0.26), compared with the correlations in Campania (0.34), Sardinia (0.38) and Lombardy (0.42).

Figure 10a: WHODAS score distributions at respective civil-invalidity cut-offs: Trentino

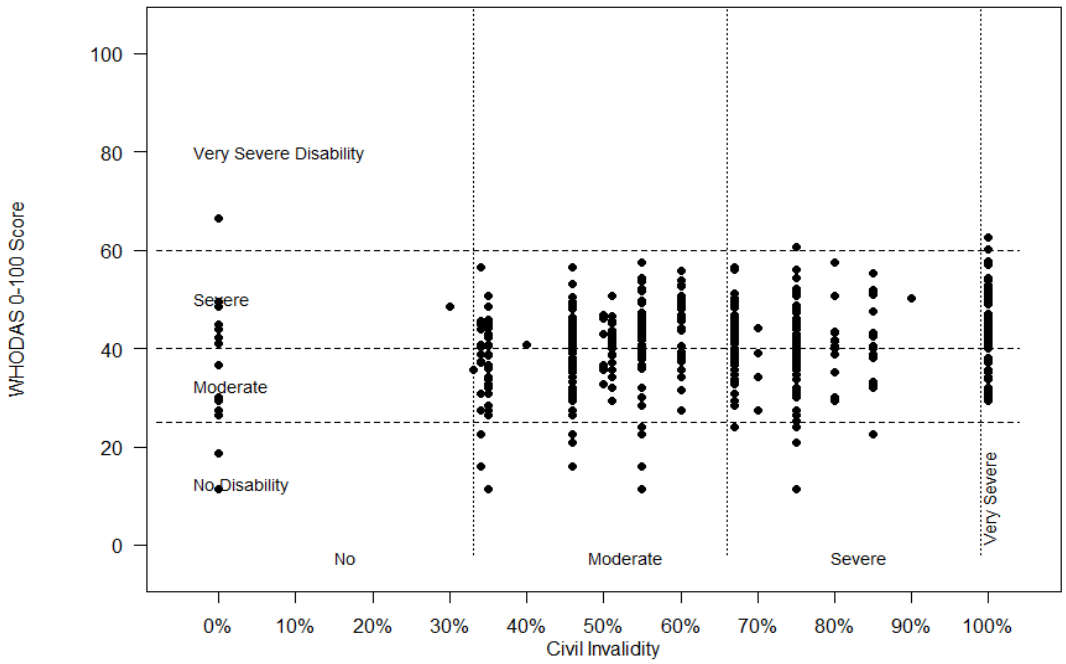


Figure 10b: WHODAS score distributions at respective civil-invalidity cut-offs: Lombardy

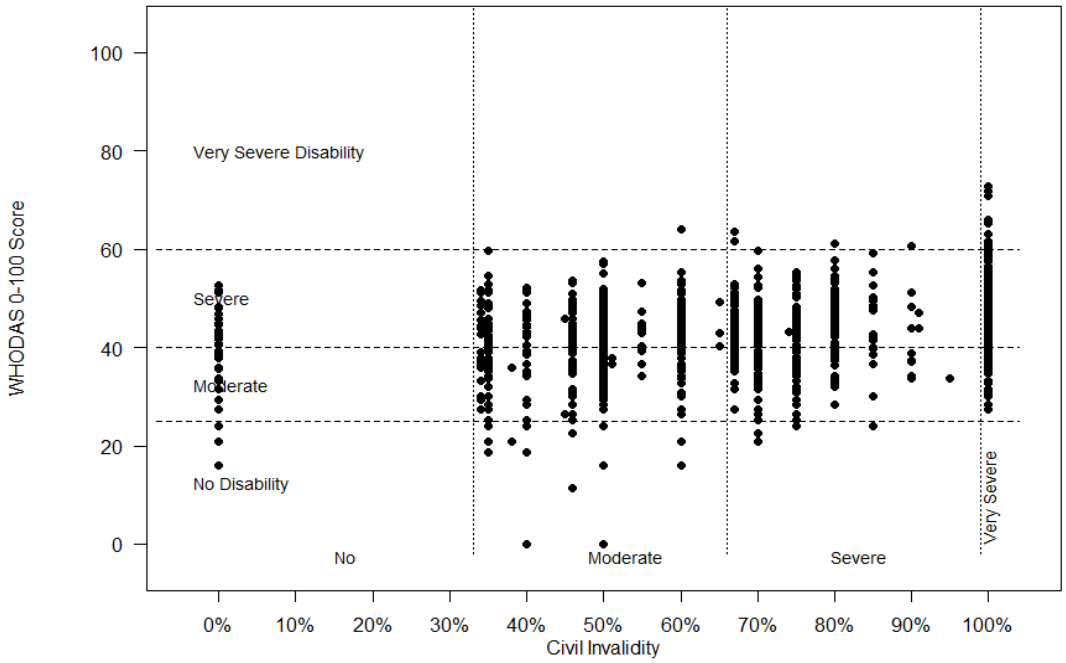


Figure 10c: WHODAS score distributions at respective civil-invalidity cut-offs: Campania

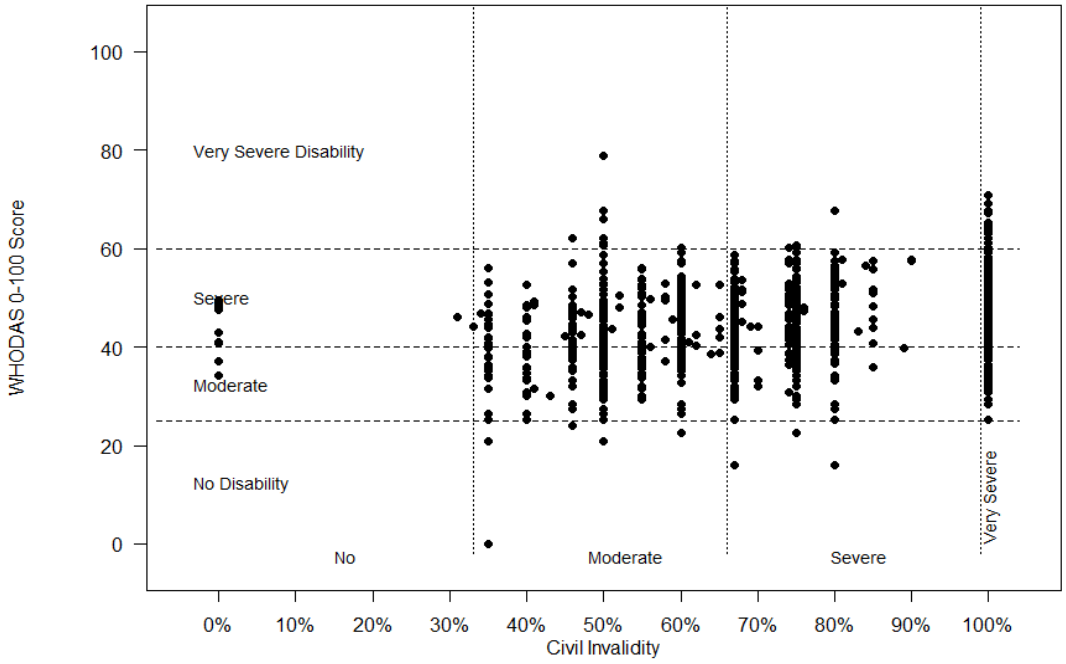
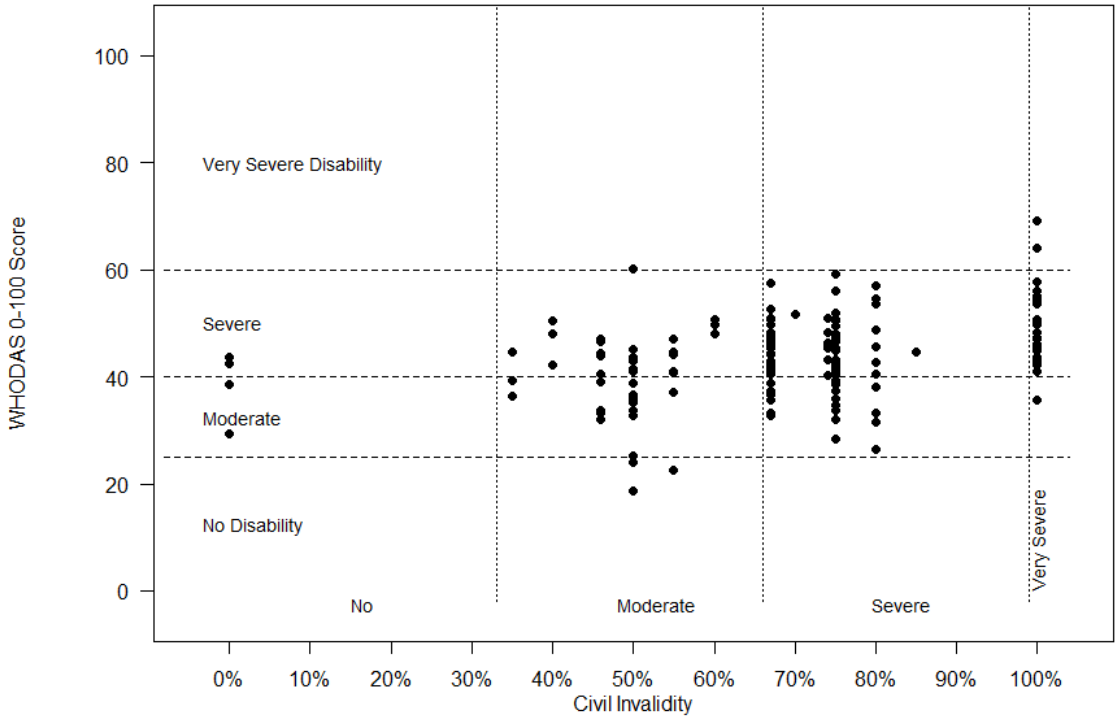


Figure 10d: WHODAS score distributions at respective civil-invalidity cut-offs: Sardinia



5. Options to include functioning elements into the assessment of civil invalidity

5.1. General considerations on the inclusion of functioning

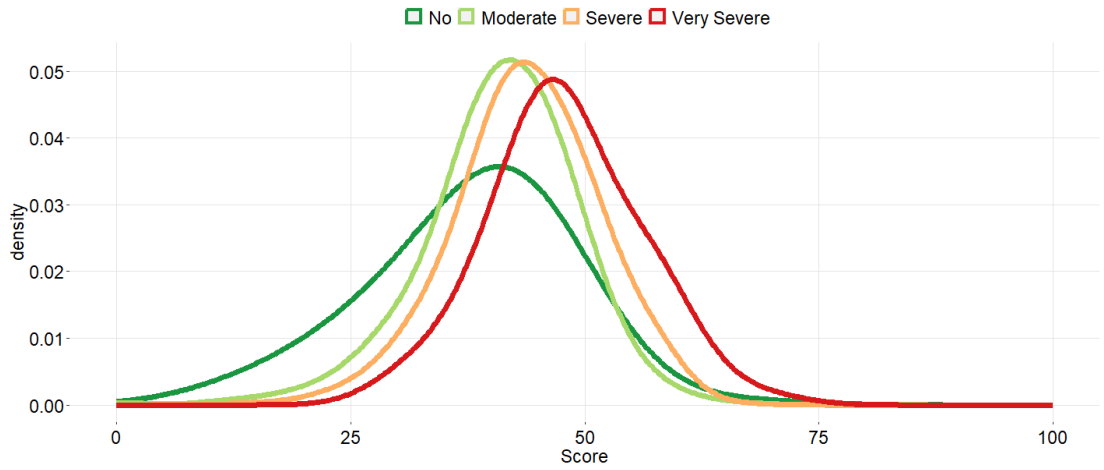
Including functioning into civil invalidity assessment improves the accuracy of the assessment and ensures that no one is left behind. One of the objectives of the analysis of the WHODAS data collected in Italy is to show that the inclusion of functioning into the current medically based assessment significantly improves its capacity to assess the experience of disability accurately and, subsequently, to allow for better assessment of the needs of people with disability. The assessment method currently in use in Italy has a strong medical focus and uses a barema-type method to translate medical information into civil invalidity percentages. The assessment is conducted in a 10-15 minute face-to-face interview between a medical board and the applicant and it is based on medical documentation, justifying the degree of impairment in diseased or injured body parts or structures. The current assessment of civil invalidity in Italy identifies invalidity degrees as percentages, with values below 33% designating mild or no disability, 33-66% moderate disability, 67-99% severe disability, and 100% very severe disability.

Barema-type methods usually involve the use of a fixed scale set out in a table according to which a certain percentage of disability is attached to specific impairments. The Barema list or table is divided into chapters covering physical or mental components of the body or the body system, and guidance is set out regarding medical benchmarks against which assessments should be made. The impairments of the person who is being assessed are compared against this list and the list automatically assigns percentage values to each impairment. Italy's civil invalidity assessment follows this process to some extent but the approach and the guidance provided to doctors differs between regions. Comparative research finds that Barema tables can be cost efficient, in that the assessment method can be based directly on a pre-existing diagnosis which only needs to be confirmed (e.g., Waddington and Priestley, 2021). In terms of the relevant disability percentage linked to a specific impairment, however, a 2002 Council of Europe report found 'no information on the reasons for choosing the levels set out in the Baremas'. There is no scienticfic proof that Barema tables assess impairments validly and reliably. Research suggests that Barema systems work best where there is no relevant 'threshold' or minimum percentage of disability which triggers entitlement to particular supports or benefits. Where such thresholds exist, medical assessors may be inclined to make an overall assessment as to whether the applicant should qualify for the support or benefit, and tailor their findings accordingly. The distribution of civil invalidity ratings in Italy very much confirms this general finding.

WHODAS functioning scores by current levels of civil invalidity demonstrate that medical assessment alone does not differentiate well between different levels of disability, also suggesting rather low reliability and precision of the civil invalidity ratings in Italy today. Figure 11 shows the density lines for the WHODAS scores for the four levels of civil invalidity. While WHODAS scores for very severe functioning restrictions stand out at least a bit (red line), the difference between severe and moderate level of civil invalidity (orange

and light green line, respectively) appears to be very small. These density lines suggest the presence of both false positives (cases with high invalidity percentage and low WHODAS score) and false negatives (cases with low invalidity percentage and high WHODAS score). Also, a more accurate assessment would show the density line of the group with no or very low level of civil invalidity (dark green line) positioned more towards the left-hand side of the figure. Again, this suggests that the medical information alone may misrepresent the true extent of individual disability experienced in daily life.

Figure 11: WHODAS-score density lines by percentage of civil invalidity (four categories)



The results presented above come as no surprise as WHODAS was designed explicitly to assess so-called whole-person disability, while the medical approach to assessing disability used in Italy does not directly assess disability but *infers* disability based on the underlying health condition or impairment. Sometimes there is a close correlation between the severity of health conditions and the severity of resulting disability; but sometimes there is no such correlation. The latter is best seen in the case of mental health problems where the impact of the person's environment may greatly increase the impact of the experience of, say, depression. This is the basic validity problem with medically based disability assessment. As pointed out above, although the presence of a health condition and associated impairment is a precondition for disability, inferring the level of disability from the presence of the underlying health condition is scientifically problematic. The level of disability that an individual experiences, as the ICF argues, is determined by the interaction between the person's health condition and associated impairments and the environment in which the person lives. WHODAS was designed to directly capture this disability experience while assessment of disability based solely on medical grounds cannot do so validly or reliably.

5.2. Real life examples

To illustrate the structural and conceptual difference between medical and functional assessment, in the following six real-life cases from the WHODAS pilot data set are presented where the civil invalidity percentage and the WHODAS score differ dramatically (see also Table 18).

- Case A is a 64 years-old man with 12 years of education, currently married and living in the community. He is self-employed. He is diagnosed with a disease of the circulatory system, an asymptomatic myocardiopathy. He has not been attributed any civil invalidity status (0%). However, based on WHODAS, he has a severe disability (score = 48), with extreme difficulties reported across all domains of daily life.

- Case B is a married man of 57 years with 8 years of education who has never been married, living in the community, and unemployed for health reasons. He has a civil invalidity of 50%, which correspond to a moderate invalidity level. He was diagnosed first with a disease of the circulatory system (hypertensive heart disease), but other conditions such as sleep apnoea, hearing loss, back problems, and anxiety are also reported. His functioning level is low, with a WHODAS score indicating a severe disability in daily life (score = 61).
- Case C is a 19-year-old man with 3 years of education. He is not married, lives in the community and is working. He has been diagnosed with a mental disorder, specifically an intellectual disability. His level of invalidity has been rated as severe with a civil invalidity percentage of 75%. In terms of functioning, his WHODAS score indicates no disability at all, with a very low score of 11.
- Case D is an 18-year-old male student in his 10th year of education. He is living in the community. He was diagnosed with a mental disorder, i.e., a mild depressive syndrome. The attributed civil invalidity is 55%, which corresponds to a moderate invalidity. The WHODAS score indicates severe difficulties in daily life with a score of 41.
- Case E is a married man of 58 years, who is working and living in the community. He is diagnosed with a neoplasm that will require surgical removal and has a poor prognosis. In addition, diabetes affecting the blood vessels and a light cardiopathy were also reported. He has been attributed a civil invalidity of 100%, indicating a very severe invalidity. Based on the WHODAS score, he experiences only moderate restrictions in his daily life (score = 29)
- Case F is a man of 57 years. He is married, living independently in the community, and unemployed for health reasons at the time of the assessment. He was diagnosed with a neoplasm with a favourable prognosis. His civil invalidity percentage has been assessed to be 0%. However, just based on the WHODAS score, his level of disability is severe (score = 66).

Table 18: Disability percentages and WHODAS scores and severity grouping – six real-life cases

Case number	Diagnosis	Civil invalidity percentage	WHODAS score
Case A	Circulatory system	0% – no invalidity	48 – severe
Case B	Circulatory system	50% – moderate	61 – very severe
Case C	Mental disorder	75% – severe	11 – no disability
Case D	Mental disorder	55% – moderate	41 – severe
Case E	Neoplasm	100% – very severe	29 – moderate
Case F	Neoplasm	0% – no invalidity	66 – very severe

The cases presented in Table 18 confirm that an assessment based on medical information alone may misrepresent the true extent of disability which an individual experiences and, in turn, fail to ensure that people can access the necessary disability supports. For example, cases A and F, with 0% civil invalidity, will not be eligible for any disability services or benefits although they experience severe and very severe difficulties in functioning, according to the WHODAS questionnaire. In contrast, case E with a civil invalidity of 100% will be eligible for personal assistance, among other things, although his disability experience in terms of functioning is moderate. More information on the lived experience of these individuals may provide explanations for the discrepancy in the disability and invalidity scores. At a minimum, a second assessment may be recommendable to prevent unfair treatment. Including functioning in disability assessment in Italy will, thus, not only improve the accuracy of assessments but will also provide more valuable input for any subsequent assessment of the actual needs of people with disability and assure a much better matching between people's actual needs and the available benefits and services. Cases A and F are not entitled to services and supports they would apparently need, while Cases C and E may be receiving or at least be entitled to services or benefits which they do not seem to need, at least not at this moment.

5.3. Options for including functioning into civil invalidity assessment in Italy

The WHODAS pilot in Italy has shown that it performs well in capturing the actual experience of disability. The question is how best to include the functioning information captured by WHODAS in the disability status assessment system in Italy. Medical information will remain relevant to disability assessment; the ICF makes it clear that without an underlying health condition and associated impairments, disability does not exist. Information about health status provides the basis for identifying specific physical and mental dimensions of activities and areas of participation vulnerable to disability, which can then be directly confirmed by the findings received from the WHODAS questionnaire. Medical information provides essential guidance on the medium and long-term trajectory of disability that the individual will experience, including whether the person faces a progressive decline in health capacity resulting in more and more disability, or the reverse, a progressive improvement. While medical information remains an essential component of disability assessment, the medical review must also change with better standardization and methodological guidelines and possibly using the ICF body functions and body structures.

As medical information is essential, this section of the report discusses possible options for combining medical and functioning information in the assessment of disability in Italy – rather than replacing the current medical approach altogether by the WHODAS questionnaire. Several methods were tested on the pilot dataset to address this question. These methods can be grouped here into three principal strategies: (1) *averaging* the medical assessment percentage with the WHODAS score to arrive at a final disability assessment score, (2) *flagging* persons whose WHODAS score and disability severity are different from the severity group based on the percentage determined by medical information alone, and (3) *scaling* the civil invalidity percentage by a certain coefficient 'x' when the WHODAS-score exceeds or falls below a certain threshold or reference value.[1] In more detail, these approaches work as follows:

(1) **Averaging** – averaging in some predetermined way the attributed disability percentage and the WHODAS score. This approach is based on the theory that, together, medical information and functioning scores contribute, to different degrees, to a realistic and valid assessment of disability.

(2) **Flagging** – identifying persons whose WHODAS severity grouping differs from the medically determined severity grouping and flagging these individuals to request from them additional information or even a full reassessment. When an individual has a WHODAS score over or below some cut-off value, this suggests that the medical score alone does not adequately capture the experience of disability and a second-level assessment should be conducted.

(3) **Scaling –** the civil invalidity percentage can be altered (i.e., raised or lowered) to reflect the WHODAS score by means of a score-based coefficient. This approach assumes that at the core of disability and civil invalidity assessment is the medical problem that the individual experiences, but at the same time, that the performance is modified (to some extent) by environmental factors that need to be understood to augment or diminish the medical score.

Averaging, flagging, and scaling are three of several potential approaches to bringing together two scores that measure different phenomena but which, together, constitute our best assessment of disability. Each approach is grounded in the ICF's understanding of disability as the outcome of an interaction between a person's underlying health condition and impairment on the one hand and the physical, human-built, interpersonal, attitudinal, social, economic, and political environment in which the person lives on the other hand. The three approaches differ, however, in how they weigh the impact of the respective medical and environmental determinants of disability. The next section describes the results of applying strategies that were tested using different weighting combinations.

[1] It is important to add that as WHODAS is used in Italy, more data are collected. This data can be analysed using the techniques from this report to continually update and recalibrate parameters and cut-off points. Moreover, the data also has other potential policy applications, including identifying disability trends and better planning for the future.

6. The impact of different policy options including functioning elements

This section presents in more detail the three options to include functioning into disability assessment in Italy. Each option follows the ICF in recommending a combination of medical and functioning assessment (with the latter provided by WHODAS). Option A is the situation in which WHODAS scores are considered, or disregarded, in a purely discretionary manner. Options B (averaging strategies), C (flagging strategies) and D (scaling strategies) are quantitative. Each option has advantages and disadvantages.

The framework for evaluating the pros and cons of every approach draws on key scientific principles that determine the credibility of any disability assessment process: validity (the extent to which the option relies on a true assessment of disability); *reliability* (the ability of the option to arrive at the same assessment of the same case by different assessors); *transparency* (the degree to which the assessment process and outcomes can be described and understood by all stakeholders); and *standardization* (the extent to which the process resists distortion or alteration over time and across locations).

6.1. Purely discretionary combination of medical and functioning components

This is the option in which an individual or committee reviews medical scores and the WHODAS scores and makes a judgment about the extent of disability as the individual or committee sees fit (Option A). This is a purely discretionary option, surprisingly common in practice. This approach is subject to manipulation, or whim, lacks validity and reliability, and is utterly non-transparent. The option is given here as a contrast to the remaining options B, C, and D, but also, in fairness, because some countries continue to rely on this option for disability assessment (strategy #1). The authors of this report do not recommend this option.[2]

6.2. Quantitative approaches including medical and functioning components

Averaging, flagging, and scaling are quantitatively driven options, very different from Option A. In different ways and for different reasons, they satisfy not only the basic psychometric assumptions of validity and reliability but each, to different degrees, strives to achieve transparency and standardization.

[2] Numerous interactions with officers involved in disability assessment in different countries suggest that medical professionals involved in the assessment of disability are confident they “know best” and can consider functioning and the experience of disability as part of the medical description of the applicant’s situation. One often hears medical assessors claim that they take functioning fully into account when examining medical records. One implicit result from the pilot is, however, that this assumption is not grounded in evidence.

Using an averaging algorithm

In the Italian pilot WHODAS data set, there is a relatively small percentage of persons indicating no functioning problems at all (only 2.3 percent), among which the majority had a moderate or severe degree of civil invalidity. Weighting the civil invalidity percentage with the WHODAS score would adjust levels of invalidity by accounting to some degree for the observed and experienced disability level assessed by the WHODAS questionnaire. To get a full sense of the range of possible approaches under Option B, four weighting schemes are shown: (i) 75% civil invalidity percentage and 25% WHODAS score; (ii) 50% each; (iii) 25% civil invalidity percentage and 75% WHODAS score; and iv) 0% civil invalidity percentage and 100% WHODAS score (represented by strategies #2 to #5). Option #5 shows the result of WHODAS alone.

Advantages of Option B: (i) An assessment of the level of functioning plays a significant role in the determination of eligibility for disability benefits so that the eligibility for benefits is not solely based on purely medical criteria. (ii) The averaging approach minimises the impact of the inherent psychometric problems with the civil invalidity percentage based on the Barema-based medical assessment. (iii) The assessment of the level of functioning is empirically and statistically verified. (iv) This option yields high levels of validity and reliability. (v) Merging the results of two assessments scaled by means of 'weighted averages' is fully objective, transparent, and non-discretionary. (vi) The method is not sample-dependent.

Disadvantages of Option B: (i) There are, potentially, an infinite number of combinations of weighting schemes (i.e., 'strategies'), each of which affects the set of eligible applicants differently and has different budgetary and political consequences. This is an unavoidable fact about the nature of disability as a continuum and the fact that there are not yet scientifically verified or objective cut-offs for severity on a 0-100 continuum. (ii) Any strategy selected will be objectionable to individuals who, under that strategy, will not be certified as disabled and thus not eligible for any benefits. This signals the need for clear and transparent information dissemination and a solid grievance redress system that may include using tools for clinical testing and determination of functioning.

Using a flagging algorithm

Six different flagging strategies are represented by strategies #6 to #11. The idea of this strategy is to highlight individuals whose civil invalidity percentage is unexpected in view of the WHODAS score. A conservative approach would be to flag individuals with scores in the upper (or lower) extremes of the WHODAS score distribution of the sample, who have a very small (or large) civil invalidity percentage (#6). The next four approaches do not use the sample distribution but the distribution of scores within civil invalidity degree groups to increase or decrease the invalidity percentage. The approach #11 combines strategies #7-10 and considers all cases that fall into one of these groups.

Advantages of Option C: (i) Scientifically robust and based on actual data. (ii) Shows that the purely medical approach to disability assessment may not accurately assess disability in many cases – in which, as reported in the WHODAS score, a person is experiencing more, or fewer, functioning problems in their lives than what the health condition is thought to imply. (iii) High levels of validity and reliability.

Disadvantages of Option C: (i) WHODAS cut-offs for different degrees of functioning problems are based on the experiences from past pilots and some evidence from the scientific literature. Sensitivity analyses are not available to this point. More precise cut-off values specific to Italy may be introduced at later time points when more information on functioning is collected (assuming WHODAS will be introduced into the existing system). (ii) Technically robust methodological and procedural instructions will have to be developed to guide the reassessment process to ensure transparency.

Even with the caveat on the cut-off points for disability severity, the flagging method may be introduced through a specifically designed two-step administrative procedure.

Using a scaling algorithm

The scaling approach, represented by strategies #12 and #13, reproduces an approach that is in some form used in some countries (e.g., Lithuania) though generally in a rather opaque way, namely, modifying the civil invalidity percentage assigned by a disability assessment committee by means of a coefficient representing functioning information (e.g., generated by a WHODAS score). The idea behind this approach

is to avoid relying on a medical determination of disability exclusively, as such an approach undervalues the actual impact of health conditions on a person's life and functioning performance.

Two strategies to illustrate the scaling approach are used (there are, in theory, many other possibilities). The first strategy would look for individuals with high disability, according to their WHODAS score, above the WHODAS cut-offs of 40 and 60 to augment their civil invalidity percentage, either by a coefficient of 1.25 (with WHODAS scores above 40) or 1.5 (with WHODAS scores above 60). Reversely, in the second strategy used, individuals with a very low disability according to their WHODAS score, below the WHODAS cut-offs of 40 and 25, are selected to reduce their civil invalidity percentage either by a coefficient of 0.95 (with WHODAS scores below 40) or 0.9 (with WHODAS scores below 25). The choice of coefficients here is to some extent driven by the objective to achieve similar impact in both directions.

Advantages of Option D: (i) Using a coefficient value generated statistically is a common and widely used approach. (ii) A coefficient approach (increasing or reducing the medically-determined civil invalidity percentage in light of the corresponding functioning score) is the most intuitive way to combine the scores of very different assessments – medical and functioning – into a single score. (iii) This option incorporates the insight that a medical determination alone can often miss instances where people have only moderate or very high disability needs. (iv) This option, because of the psychometric properties of WHODAS, would have high levels of validity and reliability.

Disadvantages of Option D: (i) As with other options, there are many possible variations of approach D with different outcomes – in this report only two possibilities are presented, as an illustrative example. Although the scaling approach itself is intuitively understandable and can be made transparent to the public, the scientific and statistical justification for Option D is therefore somewhat technical and may not be easily understandable by a lay public.

Table 19 provides an overview of the testing strategies that were considered and gives the number of individuals who would have a moderate, severe, or very severe disability after adjusting for the WHODAS score. Further, and maybe most importantly, the table also shows the number of individuals who would have their civil invalidity severity ranking changed towards a higher degree (total upshifts) or a lower degree (total downshifts). In brief, the results are as follows:

- The four *averaging* strategies show that the use of WHODAS generally generates more upshifts to higher invalidity degrees than downshifts. Giving WHODAS a weight of 25% (strategy #2) changes little, as it affects only 2.5% of the sample and of those, most would see a downshift – these are people just above one of the invalidity thresholds who seem to function well, maybe because the environment is supportive and their needs are addressed. The more weight WHODAS receives, the more people are affected and the more upshifts occur. With a 50% weight to both WHODAS and civil invalidity (strategy #3), 8.5% of the sample would be affected, with an equal number of upshifts and downshifts. With WHODAS only (strategy #5), 42% of the sample would see a change in the invalidity severity, with two-thirds seeing an upshift. Most upshifts are a shift from moderate to severe invalidity, potentially generating more eligibility for a disability allowance. On the contrary, the number of people with very severe invalidity considered to be non-self-sufficient and, thus, in need of constant care would fall drastically, from over 20% to only 2% of the sample. This suggests that current medically based disability assessment may be overestimating the degree of disability and policies may be setting the wrong priorities, and incentives.
- The six *flagging* strategies show that very few people currently receive an invalidity rating that is drastically different from their actual disability experience, as measured by WHODAS. Only 2% of the sample have extremely low or extremely high WHODAS scores (strategy #6) and only 5.5% of the sample would be flagged as having an invalidity rating very different from their WHODAS score (strategy #11). Among those 5.5%, two-thirds would potentially see a downshift in their current severity rating depending on the result of the indicated second assessment and most of them would be people classified with 100% civil invalidity although experiencing much less disability. (For supplementary flagging variants, see section 6.3).

- The coefficients chosen for the two *scaling* strategies generate a situation in which over 8% of the sample would see their invalidity rating increased because of (very) severe disability according to WHODAS (strategy #12) and, similarly, close to 8% would see their invalidity rating lowered because of no or only moderate disability experience according to WHODAS (strategy #13). The large difference in the size of the coefficients is a result of the current invalidity assessment and rating, with so many people found just above the next invalidity threshold. A clear disadvantage of strategy #12 is that it increases the already large number with a very severe invalidity rating. Combining strategy 12 and strategy 13 would imply that 16% see their rating changed.

Table 19: Overview of strategies and changes in group sizes based on the selected approaches

	#	Description	No civil invalidity	Moderate civil invalidity	Severe civil invalidity	Very severe civil invalidity	Total upshift	Total downshift
A: Discretionary	#1	Civil Invalidity cut-offs	81	1129	1076	623	0	0
B: Averaging	#2	Civil Invalidity 75%, WHODAS 25%	82	1157	1047	623	21	51
	#3	Civil Invalidity 50%, WHODAS 50%	87	1131	1068	623	116	130
	#4	Civil Invalidity 25%, WHODAS 75%	52	1059	1176	622	337	209
	#5	Civil Invalidity 0%, WHODAS 100%[1]	63	856	1921	69	768	459
C: Flagging	#6	Extreme WHODAS scores: < 24 or > 63	118	1102	1064	625	7	53
	#7	WHODAS score > 40, Civil Invalidity < 33%	40	1170	1076	623	41	0
	#8	WHODAS score > 60, Civil Invalidity < 66%	80	1120	1086	623	10	0
	#9	WHODAS score < 25, Civil Invalidity > 66%	81	1143	1062	623	0	14
	#10	WHODAS score < 40, Civil Invalidity 100%	81	1129	1170	529	0	94
	#11	Sum of approaches #7-#10	40	1174	1166	529	51	108
D: Scaling	#12	if WHODAS indicates Severe disability then Civil Invalidity x 1.25 Very severe disability then Civil Invalidity x 1.5	78	889	1125	817	243	0
	#13	if WHODAS indicates Moderate disability then Civil Invalidity x 0.95 No disability then Civil Invalidity x 0.9	105	1205	1070	529	0	218

1 This approach uses the WHODAS cut-offs: WHODAS scores < 25 indicate no disability, 20 to 40 moderate disability, 40 to 60 severe and > 60 very severe disability.

To make the different options more concrete, Table 20 illustrates them on the six real-life cases presented above (see Table 18). This illustration shows how the different strategies would potentially change the attributed level of disability for these cases, highlighting the effect of including functioning information into the current disability assessment in Italy. The table presents the expected level of disability of six selected example cases under each of the functioning inclusion strategies (with yellow = no civil invalidity; orange = moderate civil invalidity; red = severe civil invalidity; dark red = very severe civil invalidity).

Cases A and F, who with different health conditions are experiencing significant disability but have no civil invalidity degree, would see a shift in their invalidity rating under most averaging approaches but not under the scaling approach. Cases B and D, who also report more disability than identified by the civil invalidity assessment but have a moderate invalidity degree, see their invalidity rating increased by the upscaling approach and most averaging approaches. Case C, who has a severe civil invalidity rating but no disability according to WHODAS, would see the invalidity rating lowered under all approaches while Case E, with a very severe invalidity rating and a moderate level of disability, is affected by the downscaling approach and some of the averaging approaches. All cases except case D would be flagged for reassessment in at least one of the flagging approaches.

Table 20: Disability severity ranking and WHODAS scores and their integration strategies – six selected examples of somewhat extreme individual cases

			Current method severity	Averaging				Flagging*						Scaling	
	Civil Invalidity %	WHODAS score	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13
A	0	48	No invalidity						+				+		
B	50	61	Moderate					-			-		-		
C	75	11	Severe					-			-		-		
D	55	41	Moderate												
E	100	29	Very severe									-	-		
F	0	66	No Invalidity					+	+	+			+		

* the symbol "+" and "-" indicate that a case is flagged to reiterate the civil invalidity rating upwards or downwards

6.3. Reflections and conclusions

The pilot evaluation suggests that the current disability assessment system in Italy would benefit from the inclusion of functioning information into the assessment method in at least three ways:

- the assessment of disability would be more precise and accurate, reflecting the real-life experience of disability and identifying some people who are not well identified by a purely medical approach;
- the assessment would be in line with today's interdisciplinary understanding of disability to which Italy has committed already 14 years ago when it ratified the UN Convention; and
- the assessment would be harmonised with, and provide more valuable input into, any subsequent individual assessment of the actual support needs of people with disability.

The approach suggested for disability assessment is to combine medical and functioning information in some transparent form. While there are in principle many alternative methodological options for doing this, for Italy *flagging* the need for a second assessment seems to be the most meaningful and realistic way forward. This is so because the current process of civil invalidity assessment through which applicants are assigned an invalidity degree, or percentage, is strongly influenced and biased by the various thresholds in place for eligibility to various entitlements, benefits and services. Therefore, while in theory people could be assigned any percentage, in practice most applicants for a civil invalidity assessment return with a degree close to, or at, one of the critical thresholds. Technically speaking, the current assessment returns ordinally-scaled disability degrees determined by the existing thresholds rather than interval-scaled degrees that reflect the degree of the person's impairment. The consequence of this is that quantitative approaches like *scaling* or *averaging* can generate undesirable results on both ends of the spectrum. People sitting just at a threshold would easily fall below the threshold and, thus, lose critical disability entitlements, while those far away from a threshold might receive a significantly higher invalidity percentage but without any change in the types of services and benefits they are entitled to.

A related reason for the limited applicability in Italy especially of the averaging approach is the discretionary nature of Italy's civil invalidity assessment. While the assessment is intrinsically medical in nature, assessors can take people's actual situation into account if they wish: in a discretionary and untransparent way, they can increase the assigned invalidity percentage in line with any "perceived" functioning limitations – perceived, because this is done without any basis or tool to assess functioning. This problem is related to the problem that system thresholds seem to influence the assessment outcome. On the contrary, averaging would be a highly promising and adequate approach if it was used to average two independent pieces of information: the medical and the functional aspects of disability. Such a situation could be achieved also in Italy if information on these two aspects would be collected independently and the medical part of the assessment would be performed in a standardized manner with methodological guidelines applicable across the entire country.

If Italy chooses to move on with the introduction of a flagging algorithm, two aspects have to be addressed: the weight given to functioning information relative to medical information, and the structure of the entire assessment process. The first question on the relevance attached to functioning, i.e., the WHODAS score, is equal to asking how many cases "should" be flagged. Even with strategy #11, the combined result of strategies #7-#10, only about 5.5% of all applicants would be considered for a second assessment – while the remaining 94.5% would not be affected by such a reform. That is a very low share which i) does not do justice to the importance of people's actual disability experience, ii) hardly justifies a comprehensive reform, iii) would likely fail in changing everyone's mindset towards a modern view on disability and functioning and, eventually, iv) would hardly affect the adequacy and effectiveness of disability supports.

It is, therefore, useful to think about ways to increase the number of flagged cases by not only questioning and thus reassessing extreme differences between the civil invalidity percentage and the WHODAS score but also smaller differences between the medical and the functional view. For this purpose, it is useful to use the finer grid of civil invalidity thresholds, which also distinguishes lower from higher moderate invalidity and lower from higher severe invalidity, thereby creating six different invalidity categories. Similarly, the following exercise splits the moderate and severe disability groups, as measured by the WHODAS score, into two subcategories each, thereby also creating six different disability categories. The following two supplementary strategies show the range of options which Italy has.

The first supplementary strategy selects all those cases for a second assessment for which the medically-determined civil invalidity percentage on the six-category invalidity scale differs from the functionally-determined disability score on the six-category WHODAS scale. Figure 12 shows the corresponding result: cases marked in red and green are those for which the WHODAS score would imply a reassessment, with a potential downshift for the cases marked in red and an upshift for those marked in green. About one in four of the total pilot sample fall in the same category under both scales (cases marked in grey) while all others would be considered for a reassessment, with two-thirds of the flagged cases potentially considered for a downshift to a lower invalidity rating and one-third for an upshift. Most potential downshifts concern people with a 100% civil invalidity rating (very severe) or a rating between 74% and 99% (higher severe). On the contrary, most potential upshifts are people with a higher moderate invalidity rating (46%-66%).

The second supplementary strategy is less strict and allows deviations in the two scales by one category and only selects those cases for a second assessment for which the medically-determined civil invalidity percentage differs from the functionally-determined disability score by *at least two* categories. Figure 13 shows the result of this middle strategy, again marking in red and green cases with a negative or positive discrepancy between the civil invalidity rating and the WHODAS score. In about 70% of the total pilot sample, the difference between the two scales is so small that the assigned civil invalidity rating would remain untouched, while 30% would be selected for a reassessment. Of those 30%, again, about two-thirds are candidates for a potential downshift and one-third candidates for a potential upshift. In this case, most potential downshifts concern people with a 100% civil invalidity rating (very severe) while potential upshifts concern people with a lower or higher moderate invalidity rating (34%-45% or 46%-66%).

Figure 12: Flagging every case were medical invalidity and functional disability deviate by at least one category on a six-by-six civil invalidity and WHODAS disability scale

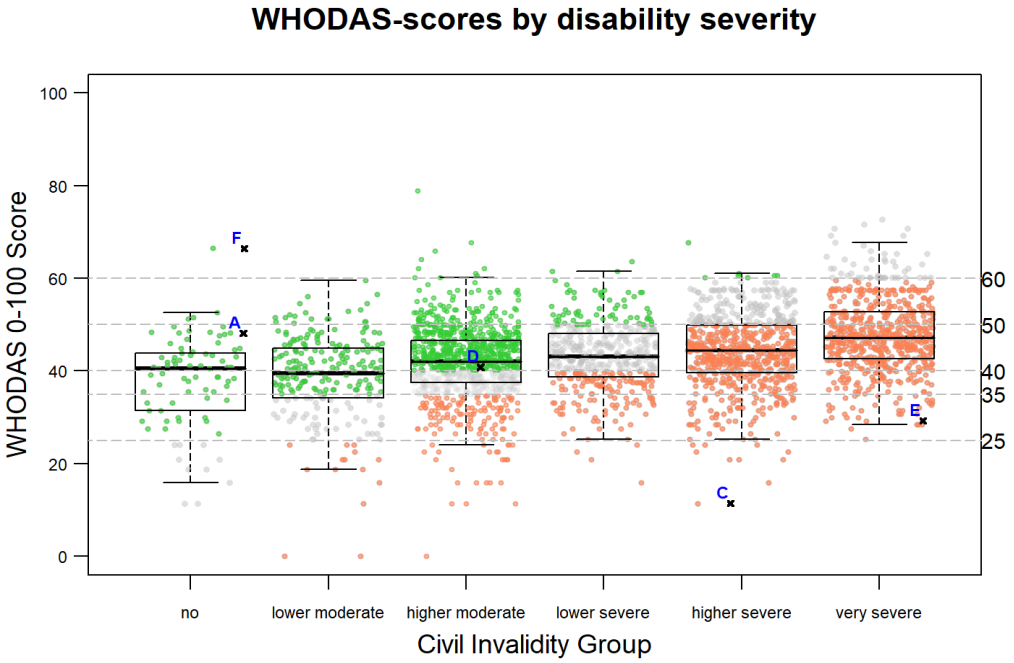
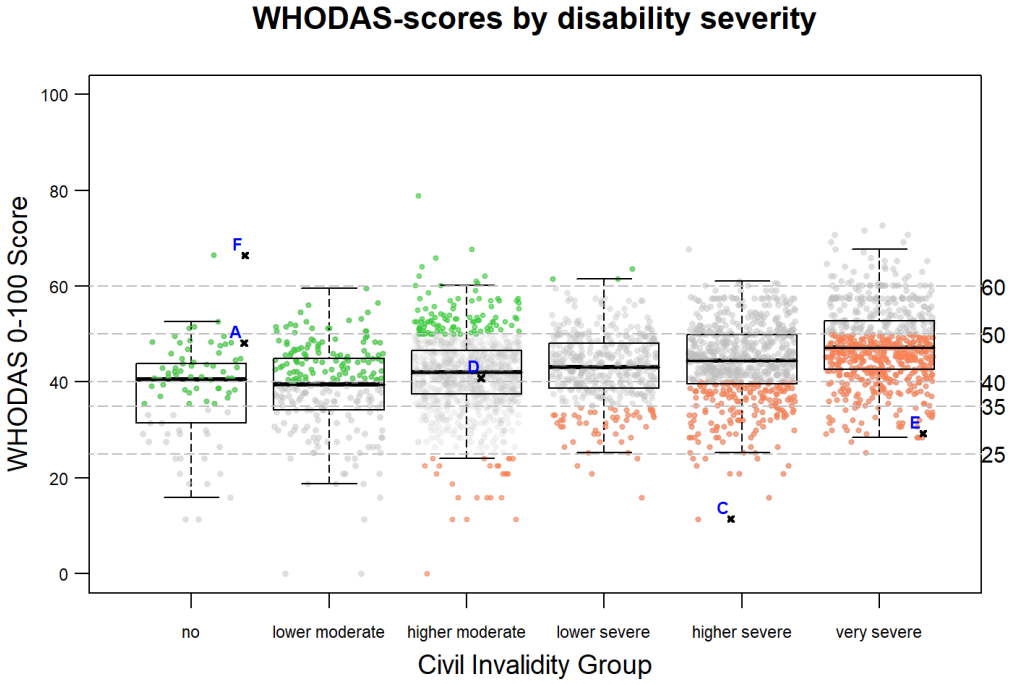


Figure 13: Flagging every case were medical invalidity and functional disability deviate by at least two categories on a six-by-six civil invalidity and WHODAS disability scale



There is no right or wrong in the choice of the particular flagging approach but the higher the importance attached to the WHODAS score, the more cases would be considered for reassessment. While these two supplementary strategies are illustrative in nature, the 30% identified in the second supplementary strategy could be a meaningful middle way for the Italian government to consider. The thresholds underlying the selection of cases for reassessment are somewhat arbitrary initially but would become more and more robust over time, as more and more data is being collected through the new assessment process.

The second aspect to consider for the introduction of a flagging algorithm is the structure of the assessment process, i.e., the question who is assessing and deciding at what stage of the process. In this context, the Italian system has a great starting advantage as the final disability rating is approved and assigned by INPS already today. This lends itself to a natural process. In a first step, medical information is assessed by the regional assessment committee, just like today, and functioning information by local social workers, as was done in the regional pilots. These two independently collected pieces of information – the person's impairment score and the person's WHODAS score – are forwarded to INPS (or any other supervisory authority) which evaluates and compares the results and decides in which cases a reassessment is needed. This is similar to today's process except that it would be done in a more transparent way and must include everyone for whom the medical and functional score deviate more than the legislation allows. If the two scores are close enough, the determination is essentially automatic and a decision on disability, by INPS, is issued. People for whom the two scores deviate are considered for a second assessment. In this case, medical assessors and social workers should sit together, examine the case and make a new joint proposal to INPS. These could be done by the medical assessors and social workers responsible for the initial evaluation, or medical assessors and social workers from INPS (or the supervisory authority).

Of course, there are additional aspects to consider within the various components. For instance, better technical and methodological guidelines would be needed for assessing doctors on how to translate impairments (via body functions and body structures) into invalidity percentages, to eliminate the current level of discretion and ensure that people with the same type and level of impairment always receive the same invalidity percentage from the assessors. Similarly, one could consider moving away from the interval scale and instead only consider groups of impairment levels, such as those used in this report.

Italy certainly has the administrative capacity to implement such a change smoothly. Italy has a cadre of experienced social workers in both the health and the social sector who could be engaged in administering WHODAS. Most Italian regions also have an advanced information system that could easily accommodate the collection and use of the information on functioning, derived from a WHODAS questionnaire, in addition to the information on the impairment. If instead of a flagging approach, which will result in a second combined medical-functional assessment in selected cases, an averaging or a scaling approach would be chosen as the method for the future, the procedure would be even easier as much of the process could be automatic. Whichever the ultimate choice might be, the result is that information on functioning will be systematically included in disability assessment using a standardized approach, and the administrative process itself will become more rigorous, standardized, and objective.

In implementing change, the Italian government will have to consider two additional, political aspects. First, any new method adopted should probably be applied to new applicants only, to make sure the change is accepted by the population. Across the OECD, only very few countries (in particular the Netherlands and the United Kingdom) have chosen to reassess current beneficiaries according to any new, reformed assessment method. Most OECD countries would, in such situations, choose to grandfather existing recipients; generally, it is considered fairer to leave existing entitlements unchanged despite the apparent inequality such an approach creates between those who were assessed before and after reform.

Second, it will be important to anticipate and manage the outcome of any reform. Whatever approach is chosen, there will be some individuals who benefit from the reform and others who will lose entitlements when compared to the current situation. As one of the conditions for reform is cost-neutrality, this issue is unavoidable. The importance given to the functioning component, relative to the medical information, will

determine the size of the two groups. Instead, Italy could also choose to produce winners only and to use functioning information only to identify people for whom the current system fails to identify their needs adequately. Such an approach would ensure that no one is left behind but would not be cost neutral.

In conclusion, this evaluation shows convincingly that the concept of disability based on functioning (via WHODAS) and the concept of civil invalidity currently in use in Italy based on impairment are hugely different. This is not surprising because one approach tries to assess the level of activity and participation and the kind and nature of problems people have in a scientifically tested way, while the other limits itself to assessing the existence, or discretionarily perceived existence, of a medical condition. The considerable difference between the two concepts demonstrates the critical importance of the inclusion of functioning into Italy's disability assessment. This will contribute to a better identification of the group of people needing support, better targeting of costly benefits and services, and a better link with regional and local needs assessments. The pilot has shown that Italy's regions are very able to implement the necessary change.

Annex A. Region specific distributions

Table A.1.: Frequencies and Percentages of WHODAS Responses: Campania

Item	No	Mild	Moderate	Severe	Extreme, cannot do	Missing
D1.1	393 (32.13%)	282 (23.06%)	299 (24.45%)	216 (17.66%)	29 (2.37%)	4 (0.33%)
D1.2	430 (35.16%)	274 (22.4%)	297 (24.28%)	189 (15.45%)	32 (2.62%)	1 (0.08%)
D1.3	410 (33.52%)	252 (20.61%)	316 (25.84%)	209 (17.09%)	34 (2.78%)	2 (0.16%)
D1.4	385 (31.48%)	258 (21.1%)	271 (22.16%)	200 (16.35%)	67 (5.48%)	42 (3.43%)
D1.5	547 (44.73%)	269 (22%)	231 (18.89%)	161 (13.16%)	14 (1.14%)	1 (0.08%)
D1.6	544 (44.48%)	235 (19.22%)	234 (19.13%)	184 (15.04%)	25 (2.04%)	1 (0.08%)
D2.1	155 (12.67%)	249 (20.36%)	284 (23.22%)	385 (31.48%)	109 (8.91%)	41 (3.35%)
D2.2	359 (29.35%)	257 (21.01%)	236 (19.3%)	335 (27.39%)	35 (2.86%)	1 (0.08%)
D2.3	443 (36.22%)	293 (23.96%)	291 (23.79%)	179 (14.64%)	17 (1.39%)	0 (0%)
D2.4	307 (25.1%)	208 (17.01%)	339 (27.72%)	292 (23.88%)	76 (6.21%)	1 (0.08%)
D2.5	101 (8.26%)	140 (11.45%)	184 (15.04%)	412 (33.69%)	254 (20.77%)	132 (10.79%)
D3.1	430 (35.16%)	276 (22.57%)	253 (20.69%)	224 (18.32%)	40 (3.27%)	0 (0%)
D3.2	436 (35.65%)	291 (23.79%)	254 (20.77%)	185 (15.13%)	20 (1.64%)	37 (3.03%)
D3.3	660 (53.97%)	315 (25.76%)	184 (15.04%)	49 (4.01%)	12 (0.98%)	3 (0.25%)
D3.4	269 (22%)	95 (7.77%)	181 (14.8%)	195 (15.94%)	97 (7.93%)	386 (31.56%)
D4.1	395 (32.3%)	253 (20.69%)	257 (21.01%)	231 (18.89%)	57 (4.66%)	30 (2.45%)
D4.2	443 (36.22%)	256 (20.93%)	262 (21.42%)	209 (17.09%)	44 (3.6%)	9 (0.74%)
D4.3	645 (52.74%)	209 (17.09%)	178 (14.55%)	133 (10.87%)	21 (1.72%)	37 (3.03%)
D4.4	308 (25.18%)	203 (16.6%)	200 (16.35%)	273 (22.32%)	162 (13.25%)	77 (6.3%)
D4.5	211 (17.25%)	132 (10.79%)	204 (16.68%)	268 (21.91%)	202 (16.52%)	206 (16.84%)
D5.1	156 (12.76%)	237 (19.38%)	346 (28.29%)	358 (29.27%)	76 (6.21%)	50 (4.09%)
D5.2	118 (9.65%)	211 (17.25%)	346 (28.29%)	363 (29.68%)	100 (8.18%)	85 (6.95%)
D5.3	116 (9.48%)	215 (17.58%)	318 (26%)	388 (31.73%)	131 (10.71%)	55 (4.5%)
D5.4	72 (5.89%)	158 (12.92%)	252 (20.61%)	411 (33.61%)	243 (19.87%)	87 (7.11%)
D5.5	57 (4.66%)	99 (8.09%)	140 (11.45%)	134 (10.96%)	88 (7.2%)	705 (57.65%)
D5.6	71 (5.81%)	92 (7.52%)	140 (11.45%)	129 (10.55%)	85 (6.95%)	706 (57.73%)
D5.7	68 (5.56%)	85 (6.95%)	134 (10.96%)	136 (11.12%)	95 (7.77%)	705 (57.65%)
D5.8	48 (3.92%)	78 (6.38%)	93 (7.6%)	169 (13.82%)	94 (7.69%)	741 (60.59%)
D6.1	169 (13.82%)	176 (14.39%)	256 (20.93%)	346 (28.29%)	225 (18.4%)	51 (4.17%)
D6.2	306 (25.02%)	256 (20.93%)	300 (24.53%)	248 (20.28%)	85 (6.95%)	28 (2.29%)
D6.3	366 (29.93%)	221 (18.07%)	272 (22.24%)	271 (22.16%)	68 (5.56%)	25 (2.04%)
D6.4	29 (2.37%)	199 (16.27%)	216 (17.66%)	562 (45.95%)	177 (14.47%)	40 (3.27%)
D6.5	25 (2.04%)	141 (11.53%)	178 (14.55%)	580 (47.42%)	293 (23.96%)	6 (0.49%)
D6.6	79 (6.46%)	241 (19.71%)	253 (20.69%)	521 (42.6%)	123 (10.06%)	6 (0.49%)
D6.7	80 (6.54%)	211 (17.25%)	294 (24.04%)	496 (40.56%)	129 (10.55%)	13 (1.06%)
D6.8	112 (9.16%)	79 (6.46%)	244 (19.95%)	399 (32.62%)	350 (28.62%)	39 (3.19%)

Table A.2.: Frequencies and Percentages of WHODAS Responses: Sardinia

Item	No	Mild	Moderate	Severe	Extreme, cannot do	Missing
D1.1	37 (20.33%)	30 (16.48%)	60 (32.97%)	42 (23.08%)	13 (7.14%)	0 (0%)
D1.2	47 (25.82%)	29 (15.93%)	59 (32.42%)	39 (21.43%)	8 (4.4%)	0 (0%)
D1.3	50 (27.47%)	26 (14.29%)	50 (27.47%)	45 (24.73%)	11 (6.04%)	0 (0%)
D1.4	60 (32.97%)	19 (10.44%)	32 (17.58%)	40 (21.98%)	20 (10.99%)	11 (6.04%)
D1.5	103 (56.59%)	26 (14.29%)	27 (14.84%)	22 (12.09%)	4 (2.2%)	0 (0%)
D1.6	105 (57.69%)	24 (13.19%)	31 (17.03%)	16 (8.79%)	6 (3.3%)	0 (0%)
D2.1	50 (27.47%)	13 (7.14%)	44 (24.18%)	46 (25.27%)	28 (15.38%)	1 (0.55%)
D2.2	49 (26.92%)	23 (12.64%)	53 (29.12%)	42 (23.08%)	15 (8.24%)	0 (0%)
D2.3	84 (46.15%)	33 (18.13%)	36 (19.78%)	24 (13.19%)	5 (2.75%)	0 (0%)
D2.4	64 (35.16%)	22 (12.09%)	42 (23.08%)	39 (21.43%)	15 (8.24%)	0 (0%)
D2.5	40 (21.98%)	14 (7.69%)	37 (20.33%)	39 (21.43%)	51 (28.02%)	1 (0.55%)
D3.1	95 (52.2%)	19 (10.44%)	30 (16.48%)	31 (17.03%)	7 (3.85%)	0 (0%)
D3.2	87 (47.8%)	29 (15.93%)	37 (20.33%)	21 (11.54%)	8 (4.4%)	0 (0%)
D3.3	135 (74.18%)	14 (7.69%)	19 (10.44%)	11 (6.04%)	3 (1.65%)	0 (0%)
D3.4	100 (54.95%)	13 (7.14%)	21 (11.54%)	19 (10.44%)	27 (14.84%)	2 (1.1%)
D4.1	107 (58.79%)	24 (13.19%)	19 (10.44%)	12 (6.59%)	18 (9.89%)	2 (1.1%)
D4.2	125 (68.68%)	17 (9.34%)	23 (12.64%)	11 (6.04%)	2 (1.1%)	4 (2.2%)
D4.3	121 (66.48%)	23 (12.64%)	18 (9.89%)	16 (8.79%)	3 (1.65%)	1 (0.55%)
D4.4	103 (56.59%)	12 (6.59%)	20 (10.99%)	13 (7.14%)	16 (8.79%)	18 (9.89%)
D4.5	101 (55.49%)	12 (6.59%)	20 (10.99%)	15 (8.24%)	28 (15.38%)	6 (3.3%)
D5.1	59 (32.42%)	27 (14.84%)	35 (19.23%)	36 (19.78%)	24 (13.19%)	1 (0.55%)
D5.2	42 (23.08%)	35 (19.23%)	35 (19.23%)	41 (22.53%)	28 (15.38%)	1 (0.55%)
D5.3	40 (21.98%)	29 (15.93%)	48 (26.37%)	33 (18.13%)	31 (17.03%)	1 (0.55%)
D5.4	24 (13.19%)	22 (12.09%)	40 (21.98%)	42 (23.08%)	52 (28.57%)	2 (1.1%)
D5.5	13 (7.14%)	8 (4.4%)	26 (14.29%)	14 (7.69%)	6 (3.3%)	115 (63.19%)
D5.6	17 (9.34%)	10 (5.49%)	19 (10.44%)	17 (9.34%)	4 (2.2%)	115 (63.19%)
D5.7	17 (9.34%)	7 (3.85%)	23 (12.64%)	14 (7.69%)	5 (2.75%)	116 (63.74%)
D5.8	12 (6.59%)	9 (4.95%)	20 (10.99%)	13 (7.14%)	12 (6.59%)	116 (63.74%)
D6.1	48 (26.37%)	26 (14.29%)	24 (13.19%)	23 (12.64%)	56 (30.77%)	5 (2.75%)
D6.2	75 (41.21%)	20 (10.99%)	36 (19.78%)	32 (17.58%)	19 (10.44%)	0 (0%)
D6.3	105 (57.69%)	20 (10.99%)	22 (12.09%)	25 (13.74%)	10 (5.49%)	0 (0%)
D6.4	17 (9.34%)	28 (15.38%)	35 (19.23%)	78 (42.86%)	24 (13.19%)	0 (0%)
D6.5	14 (7.69%)	15 (8.24%)	20 (10.99%)	81 (44.51%)	52 (28.57%)	0 (0%)
D6.6	26 (14.29%)	19 (10.44%)	38 (20.88%)	65 (35.71%)	33 (18.13%)	1 (0.55%)
D6.7	33 (18.13%)	23 (12.64%)	43 (23.63%)	57 (31.32%)	24 (13.19%)	2 (1.1%)
D6.8	36 (19.78%)	26 (14.29%)	35 (19.23%)	53 (29.12%)	32 (17.58%)	0 (0%)

Table A.3.: Frequencies and Percentages of WHODAS Responses: Lombardy

Item	No	Mild	Moderate	Severe	Extreme, cannot do	Missing
D1.1	417 (31.42%)	311 (23.44%)	346 (26.07%)	200 (15.07%)	53 (3.99%)	0 (0%)
D1.2	432 (32.55%)	342 (25.77%)	307 (23.13%)	195 (14.69%)	50 (3.77%)	1 (0.08%)
D1.3	434 (32.71%)	294 (22.16%)	336 (25.32%)	190 (14.32%)	71 (5.35%)	2 (0.15%)
D1.4	491 (37%)	267 (20.12%)	266 (20.05%)	200 (15.07%)	89 (6.71%)	14 (1.06%)
D1.5	793 (59.76%)	244 (18.39%)	189 (14.24%)	75 (5.65%)	26 (1.96%)	0 (0%)
D1.6	789 (59.46%)	242 (18.24%)	174 (13.11%)	79 (5.95%)	42 (3.17%)	1 (0.08%)
D2.1	305 (22.98%)	196 (14.77%)	303 (22.83%)	311 (23.44%)	208 (15.67%)	4 (0.3%)
D2.2	386 (29.09%)	269 (20.27%)	329 (24.79%)	243 (18.31%)	98 (7.39%)	2 (0.15%)
D2.3	583 (43.93%)	312 (23.51%)	299 (22.53%)	98 (7.39%)	33 (2.49%)	2 (0.15%)
D2.4	451 (33.99%)	225 (16.96%)	329 (24.79%)	213 (16.05%)	107 (8.06%)	2 (0.15%)
D2.5	266 (20.05%)	183 (13.79%)	235 (17.71%)	254 (19.14%)	347 (26.15%)	42 (3.17%)
D3.1	712 (53.65%)	214 (16.13%)	244 (18.39%)	118 (8.89%)	38 (2.86%)	1 (0.08%)
D3.2	677 (51.02%)	282 (21.25%)	225 (16.96%)	112 (8.44%)	30 (2.26%)	1 (0.08%)
D3.3	909 (68.5%)	199 (15%)	130 (9.8%)	68 (5.12%)	17 (1.28%)	4 (0.3%)
D3.4	556 (41.9%)	131 (9.87%)	135 (10.17%)	88 (6.63%)	97 (7.31%)	320 (24.11%)
D4.1	695 (52.37%)	250 (18.84%)	209 (15.75%)	116 (8.74%)	47 (3.54%)	10 (0.75%)
D4.2	801 (60.36%)	222 (16.73%)	153 (11.53%)	102 (7.69%)	40 (3.01%)	9 (0.68%)
D4.3	867 (65.34%)	216 (16.28%)	152 (11.45%)	66 (4.97%)	21 (1.58%)	5 (0.38%)
D4.4	600 (45.21%)	176 (13.26%)	197 (14.85%)	137 (10.32%)	112 (8.44%)	105 (7.91%)
D4.5	400 (30.14%)	139 (10.47%)	167 (12.58%)	149 (11.23%)	192 (14.47%)	280 (21.1%)
D5.1	57 (4.3%)	42 (3.17%)	63 (4.75%)	36 (2.71%)	14 (1.06%)	1115 (84.02%)
D5.2	63 (4.75%)	35 (2.64%)	61 (4.6%)	39 (2.94%)	15 (1.13%)	1114 (83.95%)
D5.3	254 (19.14%)	228 (17.18%)	379 (28.56%)	297 (22.38%)	153 (11.53%)	16 (1.21%)
D5.4	165 (12.43%)	192 (14.47%)	370 (27.88%)	347 (26.15%)	236 (17.78%)	17 (1.28%)
D5.5	96 (7.23%)	139 (10.47%)	238 (17.94%)	173 (13.04%)	99 (7.46%)	582 (43.86%)
D5.6	159 (11.98%)	143 (10.78%)	196 (14.77%)	154 (11.61%)	94 (7.08%)	581 (43.78%)
D5.7	153 (11.53%)	136 (10.25%)	204 (15.37%)	150 (11.3%)	98 (7.39%)	586 (44.16%)
D5.8	103 (7.76%)	109 (8.21%)	213 (16.05%)	180 (13.56%)	136 (10.25%)	586 (44.16%)
D6.1	354 (26.68%)	179 (13.49%)	264 (19.89%)	236 (17.78%)	237 (17.86%)	57 (4.3%)
D6.2	566 (42.65%)	215 (16.2%)	244 (18.39%)	193 (14.54%)	98 (7.39%)	11 (0.83%)
D6.3	616 (46.42%)	229 (17.26%)	208 (15.67%)	172 (12.96%)	92 (6.93%)	10 (0.75%)
D6.4	78 (5.88%)	168 (12.66%)	326 (24.57%)	513 (38.66%)	239 (18.01%)	3 (0.23%)
D6.5	49 (3.69%)	109 (8.21%)	234 (17.63%)	483 (36.4%)	446 (33.61%)	6 (0.45%)
D6.6	193 (14.54%)	253 (19.07%)	336 (25.32%)	362 (27.28%)	177 (13.34%)	6 (0.45%)
D6.7	206 (15.52%)	211 (15.9%)	345 (26%)	388 (29.24%)	152 (11.45%)	25 (1.88%)
D6.8	296 (22.31%)	228 (17.18%)	287 (21.63%)	267 (20.12%)	208 (15.67%)	41 (3.09%)

Table A.4.: Frequencies and Percentages of WHODAS Responses: Trentino

Item	No	Mild	Moderate	Severe	Extreme, cannot do	Missing
D1.1	186 (36.47%)	100 (19.61%)	140 (27.45%)	70 (13.73%)	12 (2.35%)	2 (0.39%)
D1.2	189 (37.06%)	114 (22.35%)	122 (23.92%)	74 (14.51%)	10 (1.96%)	1 (0.2%)
D1.3	175 (34.31%)	139 (27.25%)	108 (21.18%)	66 (12.94%)	18 (3.53%)	4 (0.78%)
D1.4	196 (38.43%)	110 (21.57%)	94 (18.43%)	71 (13.92%)	19 (3.73%)	20 (3.92%)
D1.5	263 (51.57%)	126 (24.71%)	71 (13.92%)	40 (7.84%)	9 (1.76%)	1 (0.2%)
D1.6	277 (54.31%)	104 (20.39%)	79 (15.49%)	40 (7.84%)	10 (1.96%)	0 (0%)
D2.1	128 (25.1%)	77 (15.1%)	103 (20.2%)	123 (24.12%)	78 (15.29%)	1 (0.2%)
D2.2	154 (30.2%)	103 (20.2%)	128 (25.1%)	97 (19.02%)	27 (5.29%)	1 (0.2%)
D2.3	258 (50.59%)	112 (21.96%)	101 (19.8%)	33 (6.47%)	6 (1.18%)	0 (0%)
D2.4	214 (41.96%)	106 (20.78%)	97 (19.02%)	65 (12.75%)	27 (5.29%)	1 (0.2%)
D2.5	131 (25.69%)	75 (14.71%)	101 (19.8%)	98 (19.22%)	99 (19.41%)	6 (1.18%)
D3.1	292 (57.25%)	95 (18.63%)	72 (14.12%)	38 (7.45%)	13 (2.55%)	0 (0%)
D3.2	269 (52.75%)	106 (20.78%)	85 (16.67%)	41 (8.04%)	9 (1.76%)	0 (0%)
D3.3	372 (72.94%)	68 (13.33%)	42 (8.24%)	23 (4.51%)	5 (0.98%)	0 (0%)
D3.4	280 (54.9%)	50 (9.8%)	32 (6.27%)	25 (4.9%)	45 (8.82%)	78 (15.29%)
D4.1	273 (53.53%)	102 (20%)	53 (10.39%)	53 (10.39%)	16 (3.14%)	13 (2.55%)
D4.2	305 (59.8%)	74 (14.51%)	64 (12.55%)	39 (7.65%)	16 (3.14%)	12 (2.35%)
D4.3	331 (64.9%)	83 (16.27%)	68 (13.33%)	22 (4.31%)	4 (0.78%)	2 (0.39%)
D4.4	244 (47.84%)	59 (11.57%)	54 (10.59%)	58 (11.37%)	37 (7.25%)	58 (11.37%)
D4.5	205 (40.2%)	60 (11.76%)	58 (11.37%)	48 (9.41%)	35 (6.86%)	104 (20.39%)
D5.1	163 (31.96%)	107 (20.98%)	116 (22.75%)	90 (17.65%)	28 (5.49%)	6 (1.18%)
D5.2	159 (31.18%)	110 (21.57%)	114 (22.35%)	83 (16.27%)	34 (6.67%)	10 (1.96%)
D5.3	152 (29.8%)	96 (18.82%)	125 (24.51%)	89 (17.45%)	39 (7.65%)	9 (1.76%)
D5.4	104 (20.39%)	87 (17.06%)	127 (24.9%)	124 (24.31%)	57 (11.18%)	11 (2.16%)
D5.5	75 (14.71%)	44 (8.63%)	78 (15.29%)	79 (15.49%)	20 (3.92%)	214 (41.96%)
D5.6	102 (20%)	57 (11.18%)	58 (11.37%)	63 (12.35%)	15 (2.94%)	215 (42.16%)
D5.7	96 (18.82%)	51 (10%)	69 (13.53%)	59 (11.57%)	20 (3.92%)	215 (42.16%)
D5.8	69 (13.53%)	54 (10.59%)	70 (13.73%)	70 (13.73%)	32 (6.27%)	215 (42.16%)
D6.1	182 (35.69%)	50 (9.8%)	66 (12.94%)	63 (12.35%)	79 (15.49%)	70 (13.73%)
D6.2	206 (40.39%)	85 (16.67%)	85 (16.67%)	89 (17.45%)	29 (5.69%)	16 (3.14%)
D6.3	229 (44.9%)	97 (19.02%)	71 (13.92%)	75 (14.71%)	31 (6.08%)	7 (1.37%)
D6.4	34 (6.67%)	107 (20.98%)	105 (20.59%)	204 (40%)	58 (11.37%)	2 (0.39%)
D6.5	35 (6.86%)	53 (10.39%)	91 (17.84%)	213 (41.76%)	117 (22.94%)	1 (0.2%)
D6.6	102 (20%)	100 (19.61%)	118 (23.14%)	129 (25.29%)	55 (10.78%)	6 (1.18%)
D6.7	116 (22.75%)	95 (18.63%)	94 (18.43%)	145 (28.43%)	43 (8.43%)	17 (3.33%)
D6.8	153 (30%)	100 (19.61%)	93 (18.24%)	98 (19.22%)	50 (9.8%)	16 (3.14%)

Annex B. Illustration of inclusion strategies

What follows, illustrates graphically how the averaging and flagging options function. The averaging strategy is illustrated by means of five relative weightings of the disability percentage and the WHODAS score – as in the averaging strategies discussed in this report. The averaging approach can be easily depicted by means of a cartesian coordinate system with the disability percentage on the x-axis and the WHODAS score on the y-axis. The two weighted cut-off lines separate between no and moderate invalidity and between moderate and severe civil invalidity levels. Like a clock hand, the separation line moves with increasing weight of the WHODAS score, with individuals who are either shifted upward for a higher disability level or downwards with lesser functioning problems or disability. The coordinate system approach could be implemented in practice to actually 'locate' specific individuals on the graph, based on their disability percentage and their WHODAS score. This makes it possible, at a glance, to see if the scores of an individual are congruent or if there is a discrepancy between the ratings that would require a reiteration or reassessment of the case and finally a re-attribution of a disability percentage.

The flagging strategy is illustrated by means of five figures with boxplots. Separate boxplots are drawn for each civil invalidity level. The WHODAS scores of persons are located in the boxplots corresponding to their civil invalidity rating. If the WHODAS scores cross a certain cut-off, a person can be flagged and selected for reconsideration of their invalidity assessment.

For concreteness, the six described individuals (A, B, C, D, E, and F) are also located in each graph.

Actual Approach

Strategy #1 (Figure B.1) only considers the civil invalidity percentage, the cartesian field is divided vertically at a cut-off of 33% and 66%, with individuals having no invalidity in the yellow field, individuals with moderate invalidity in the orange field, and individuals with a severe to very severe invalidity in the red field. This colouring scheme will be used in all Figures illustrating the averaging approach.

STRATEGY #1
(Civil Invalidity 100% and WHODAS 0%)

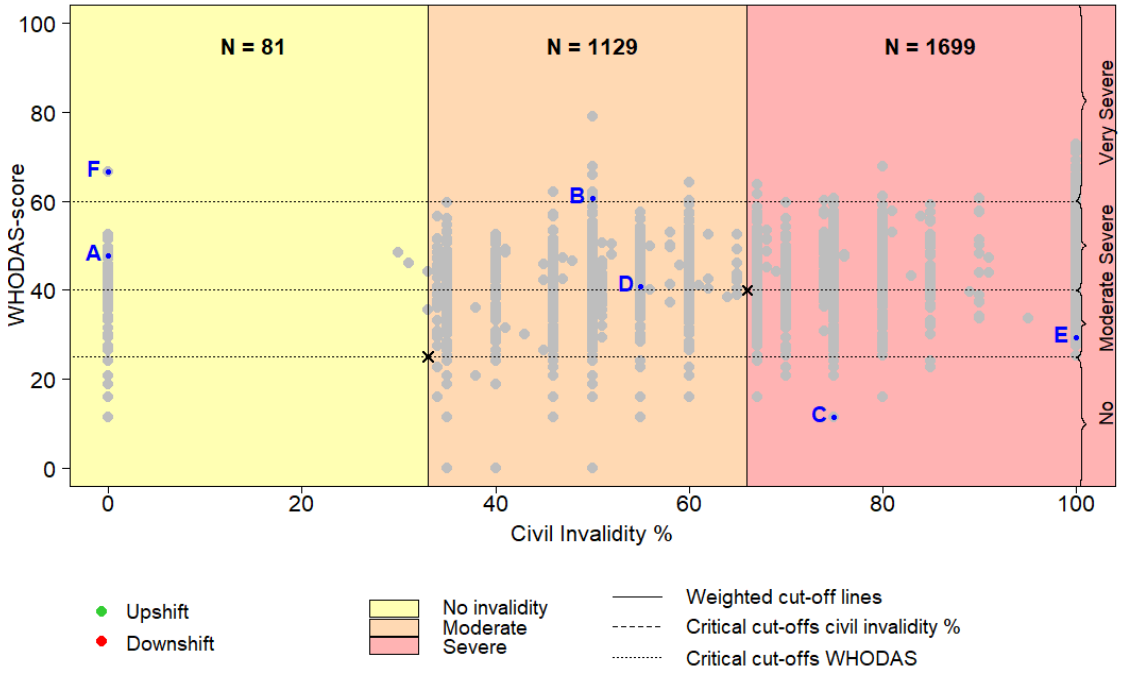


Figure B.1: Civil Invalidity percentage 100% and WHODAS score 0%

Without any adjustment to the actual approach, cases A and F have no invalidity, cases B and D have a moderate invalidity rating, and cases C and E a severe and very severe invalidity rating.

Illustrating Averaging

In the averaging strategies #2 to #5, the cut-off lines are gradually rotated around two cut-off points for the WHODAS scores that separate no from moderate functioning problems (WHODAS-score = 25) and moderate from severe functioning problems (WHODAS-score = 40).

STRATEGY #2
(Civil Invalidity 75% and WHODAS 25%)

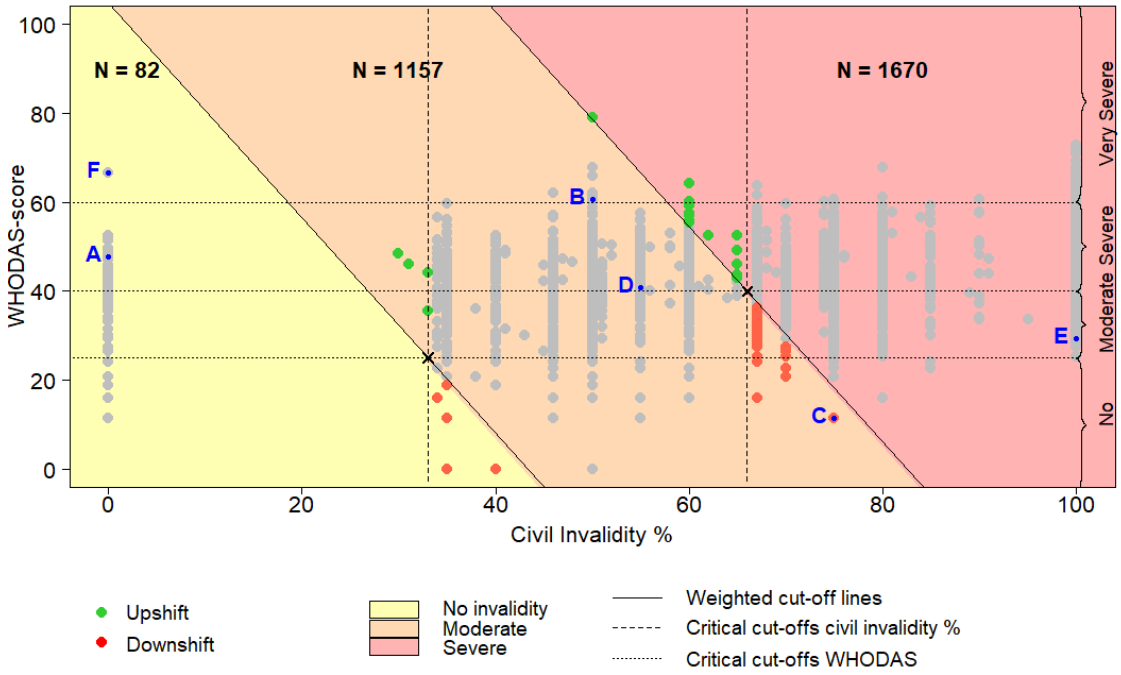


Figure B.2: Civil Invalidity percentage 75% and WHODAS score 25%

In strategy #2, WHODAS contributes 25% to the disability assessment. This would change the disability level of Case C who would shift from severe to a moderate invalidity level group. Case C, is a young man of 19 years with an intellectual disability, living and working in the community. His invalidity has been rated as severe while, based on the assessment with WHODAS, he would not be having functioning problems. Applying this strategy, would, at the level of the survey population upshift N = 21 individuals towards more disability and downshift N = 51 individuals towards less disability.

STRATEGY #3
(Civil Invalidity 50% and WHODAS 50%)

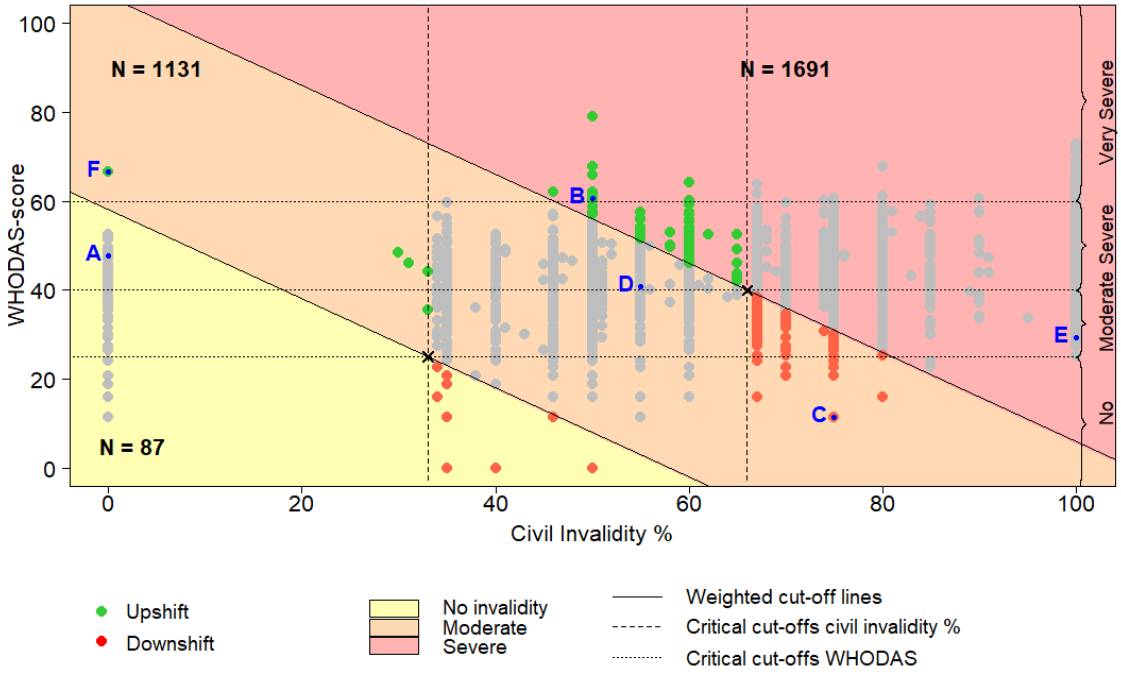


Figure B.3: Civil Invalidity percentage 50% and WHODAS score 50%

In strategy #3, WHODAS contributes 50%. This would now additionally affect the invalidity rating of Case B and Case F. Case B is a married man of 57 years, who is unemployed for health reasons. He has been reported as having a disease of the circulatory system, and his invalidity is rated as 50% (i.e., moderate), however, based on WHODAS his functioning in day-to-day life is severely restricted. Case F is also a man of 57 years, also unemployed for health reason. He has been diagnosed with a neoplasm with favourable diagnosis. While he is not eligible for any benefits (civil invalidity is 0%), he reports severe disability in the WHODAS assessment. Both cases, would have to be reconsidered, to make sense of the high levels of disability reported by these two persons in the light of the low ratings of civil invalidity. In total, after inclusion of the WHODAS ratings, the level of civil invalidity could be considered too low for 116 individuals and too high for 130 individuals in the pilot sample.

STRATEGY #4
(Civil Invalidity 25% and WHODAS 75%)

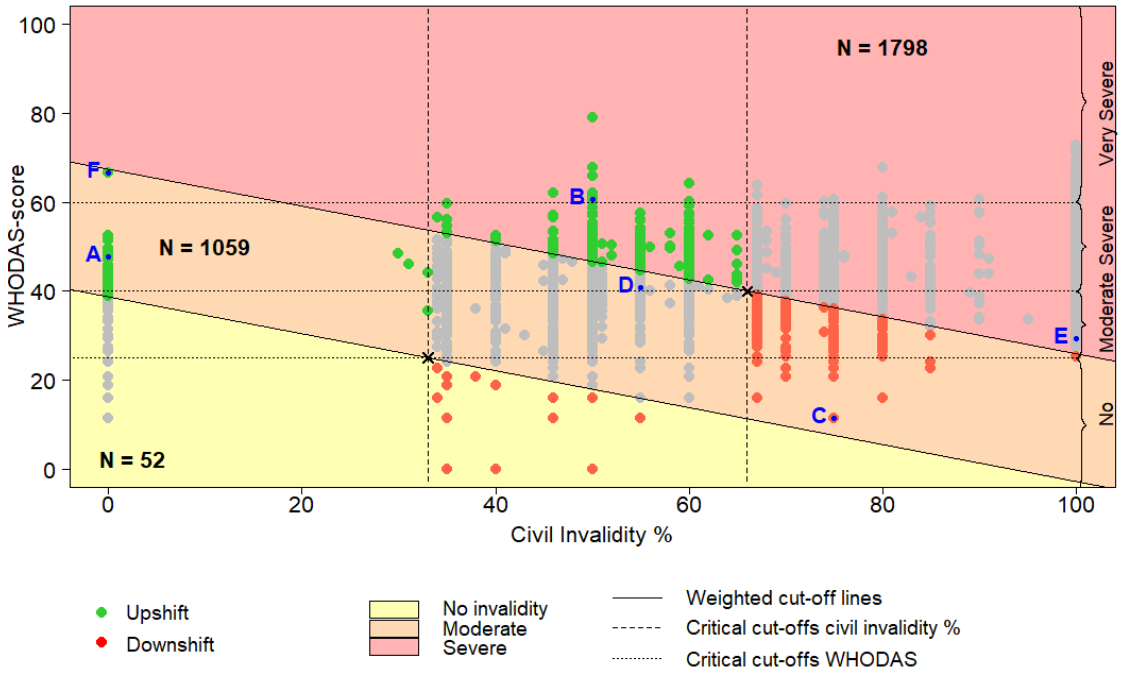


Figure B.4: Civil Invalidity percentage 25% and WHODAS score 75%

In strategy #4, the functioning assessment would receive more weight, with WHODAS contributing 75% to the disability assessment. In this strategy, Case A would be reconsidered, given his relatively high WHODAS score considering that the civil invalidity percentage was estimated to be 0%. In total, after reweighting the civil invalidity percentage by means of the WHODAS ratings, levels of civil invalidity would be seen as too low for 337 individuals and too high for 459 individuals in the pilot sample.

STRATEGY #5
(Civil Invalidity 0% and WHODAS 100%)

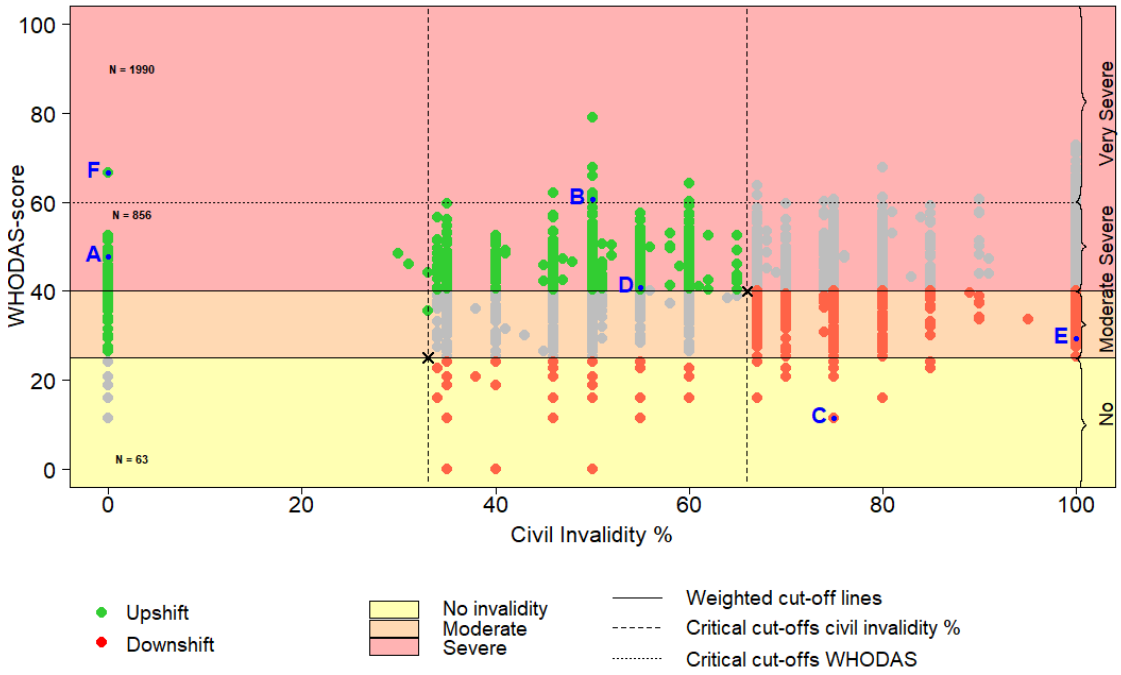


Figure B.5: Civil Invalidity percentage 0% and WHODAS score 100%

Strategy #5 is the most extreme strategy as it would only account for the functioning information derived based on the WHODAS score. Case A and F, with a civil invalidity percentage of 0% would need to be rediscussed, as the WHODAS assessment supports that both face severe functioning problems in daily life. The degree of invalidity of Cases B and D would also need to be reconsidered as the severity indicated by WHODAS is higher than what was found with the civil invalidity percentage. Case C on the other hand, appears to have very little disability based on the WHODAS score, although his civil invalidity rating is high. This could be questioned. When functioning enters the equation, the level of civil invalidity would be seen as too low for 768 individuals and too high for 459 individuals in the full pilot sample.

Illustrating Flagging

Flagging is a strategy that is less quantitatively driven and mainly consists in highlighting individuals with WHODAS-scores above or below a critical cut-off value. In what follows, several flagging approaches are briefly discussed, also in light of the 6 individual cases.

STRATEGY #6: Flagging extreme WHODAS scores

Figure B.6: Extreme WHODAS scores

Strategy #6 locates the extremes of the WHODAS-score distribution, to especially flag individuals located in the lower extreme (<24.1, i.e., no functioning restrictions) and having a high civil invalidity percentage or, reversely, individuals located in the upper extreme (>63.2, i.e., very severe functioning restrictions) and having a low civil invalidity rating. With this approach, the civil invalidity of 53 individuals is expected too high, given their good functioning (red dots). On the other hand, the civil invalidity of 7 individuals can be expected as too low, given the high level of disability measured by WHODAS (green dots). With this approach, the low WHODAS scores and severe civil invalidity rating of Cases B, C, and F would be flagged.

STRATEGY #7 Having a WHODAS score > 40 and a civil invalidity percentage < 33%

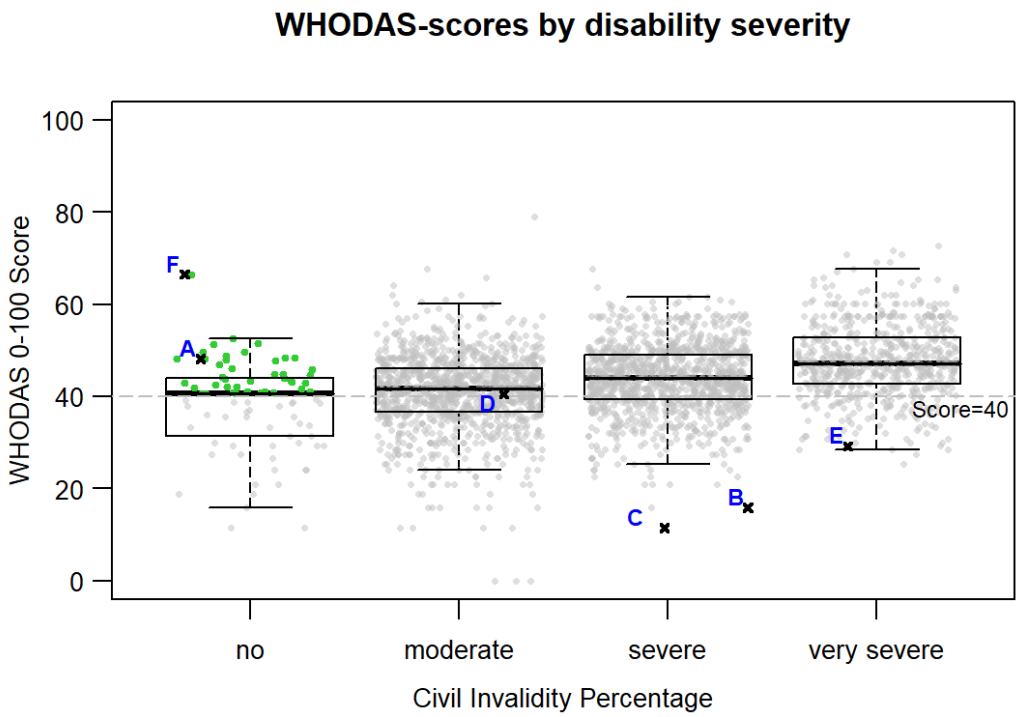


Figure B.7: Severe or very severe disability based on the WHODAS score

Strategy #7 flags and may benefit individuals that have a WHODAS score above 40, i.e., individuals with severe to very severe functioning problems, but who have been attributed a civil invalidity percentage below 33%. For these persons, here the green dots should be reconsidered, their invalidity percentage may not entirely capture the functioning problems that they experience in daily life. With this strategy, Cases A and F may shift upwards and obtain benefits according to a higher civil invalidity level. Out of te full pilot sample, with this strategy N = 41 individuals would be identified as having functioning restrictions above expectation.

STRATEGY #8: Having a WHODAS score > 60 and a civil invalidity percentage < 66%

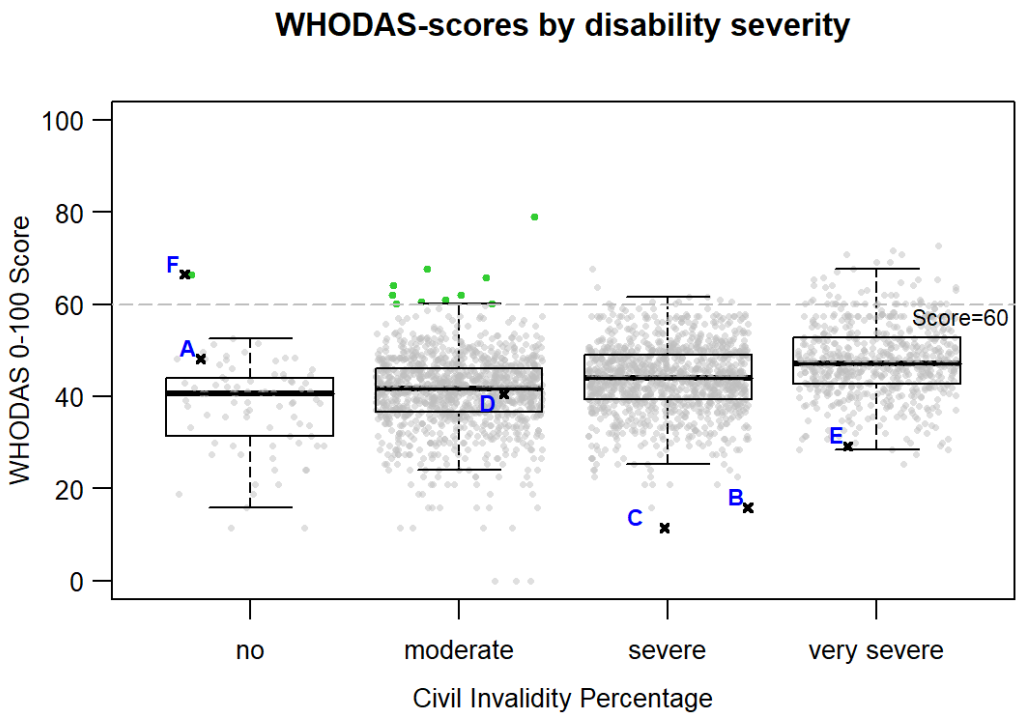


Figure B.8: Very severe disability based on the WHODAS score

Strategy #8 searches for individuals having a disability score indicating very severe functioning restriction, i.e., a WHODAS score > 60 (very severe), and a civil invalidity of < 66% (no to moderate invalidity). With this approach, 11 individuals would have been flagged for a second assessment. Only Case F would be retained, given that the disability cut-off is very high.

STRATEGY #9 Having a WHODAS score < 25 and a civil invalidity > 66%

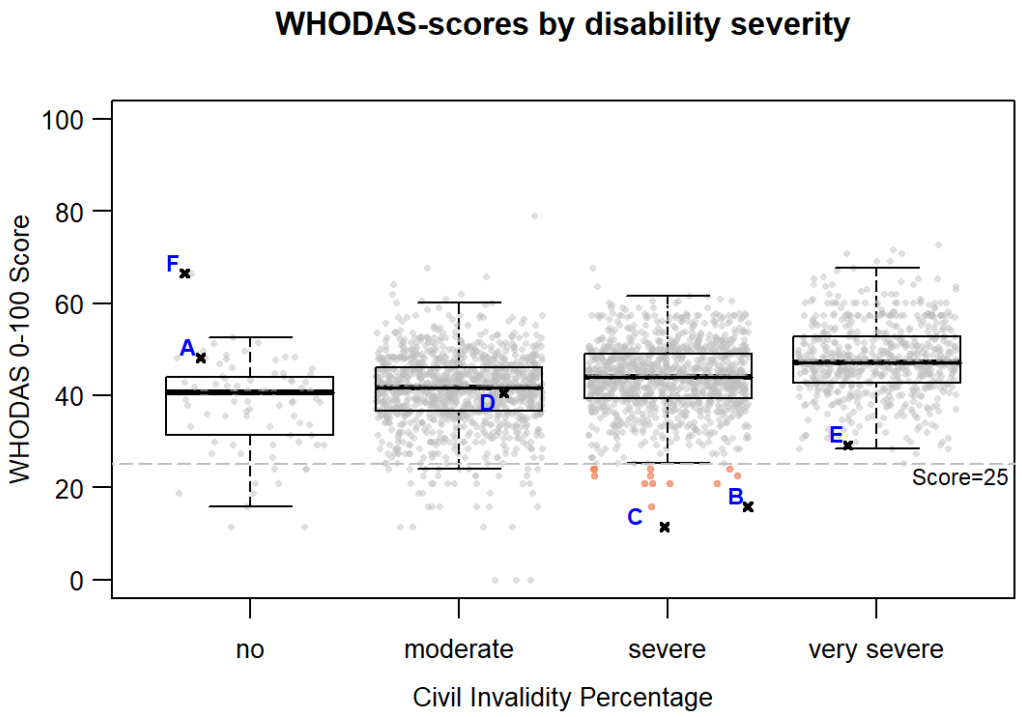


Figure B.9: No disability based on the WHODAS score

The strategy #9 and also strategy #10 that follows could be combined with one of the two previous approaches (#7 or #8). Strategy #9 aims to detect cases with very low WHODAS scores in groups of severe to very severe degrees of civil invalidity. Here the chosen WHODAS cut-off is 25, which indicates no disability or no functioning restrictions. With this approach, a total of 14 individuals would have their civil invalidity reconsidered to understand their high civil invalidity ratings.

STRATEGY #10 Having a WHODAS score < 40 and a civil invalidity percentage of 100%

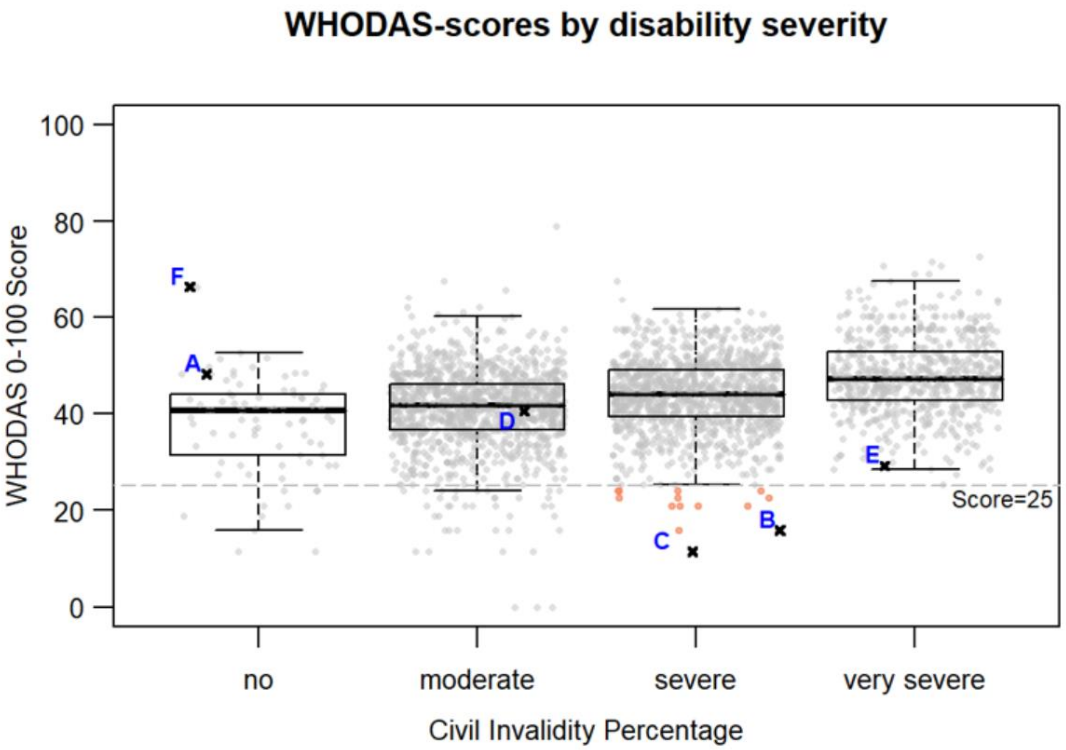


Figure B.10: Moderate disability based on the WHODAS score

Strategy #10 aims to detect cases with lower WHODAS scores in the population with very severe degrees of civil invalidity (100%). Here the chosen WHODAS cut-off is 40, which indicates no to moderate disability. With this approach a total of 94 individuals with the highest civil invalidity rating and only moderate disability in daily life would be flagged for a reconsideration or a justification of the civil invalidity rating.

References

Bond, Trevor G., and Christine M. Fox. 2001. *Applying the Rasch Model : Fundamental Measurement in the Human Sciences*. Mahwah, NJ: L. Erlbaum.

Boon, Wiiliam, John Staver and Melissa Yale. 2014. *Rasch Analysis in the Human Sciences*. Dordrecht: Springer Science+Business Media.

Federici, Stefano, Marco Bracalenti, Fabio Meloni, and Juan V. Luciano. 2017. "World Health Organization Disability Assessment Schedule 2.0: An International Systematic Review." *Disability and Rehabilitation* 39 (23): 2347–80. https://doi.org/10.1080/09638288.2016.1223177.

Fellinghauer, Carolina Saskia, Birgit Prodinger, and Alan Tennant. 2018. "The Impact of Missing Values and Single Imputation Upon Rasch Analysis Outcomes: A Simulation Study." *Journal of Applied Measurement* 19 (1): 1–25.

Ferrer, Michele Lacerda Pereira, Monica Rodrigues Perracini, Flávio Rebustini, and Cassia Maria Buchalla. 2019. "WHODAS 2.0-BO: Normative Data for the Assessment of Disability in Older Adults." *Revista de Saude Publica* 53: 19. https://doi.org/10.11606/S1518-8787.2019053000586.

Holland, P. W., and H. Wainer. 1993. *Differential Item Functioning*. Edited by N. J. Hillsdale. Erlbaum.

Masters, Geoff N. 1982. "A Rasch Model for Partial Credit Scoring." *Psychometrika* 47 (June): 149–74.

Mayrink, Jussara, Renato T. Souza, Carla Silveira, José P. Guida, Maria L. Costa, Mary A. Parpinelli, Rodolfo C. Pacagnella, et al. 2018. "Reference Ranges of the WHO Disability Assessment Schedule (WHODAS 2.0) Score and Diagnostic Validity of Its 12-Item Version in Identifying Altered Functioning in Healthy Postpartum Women." *International Journal of Gynecology & Obstetrics* 141 (S1): 48–54. https://doi.org/https://doi.org/10.1002/ijgo.12466.

Nunnally, Jum C., and Ira H. Bernstein. 1994. *Psychometric Theory*. 3rd ed. New York ; London: McGraw-Hill.

Rasch, G. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: [s.n.].

Smith, E. V. 2002. "Detecting and Evaluating the Impact of Multidimensionality Using Item Fit Statistics and Principal Component Analysis of Residuals." *J Appl Meas* 3 (2): 205–31.

Smith, R. M., and C. Y. Miao. 1994. "Assessing Unidimensionality for Rasch Measurement." In *Objective Measurement: Theory into Practice. Volume 2*. Greenwich: Ablex: M. Wilson.

Smith, R. M., R. E. Schumacker, and M. J. Bush. 1998. "Using Item Mean Squares to Evaluate Fit to the Rasch Model." *J Outcome Meas* 2 (1): 66–78.

Stekhoven, D. J., and P. Buhlmann. 2012. "MissForest–Non-Parametric Missing Value Imputation for Mixed-Type Data." *Bioinformatics* 28 (1): 112–18. https://doi.org/10.1093/bioinformatics/btr597.

Team, R Core. 2016. "R: A Language and Environment for Statistical Computing."

Tennant, A., and P. G. Conaghan. 2007. "The Rasch Measurement Model in Rheumatology: What Is It and Why Use It? When Should It Be Applied, and What Should One Look for in a Rasch Paper?" *Arthritis Rheum* 57 (8): 1358–62. https://doi.org/10.1002/art.23108.

Waddington, L., & Priestley, M. (2021). A human rights approach to disability assessment. Journal of International and Comparative Social Policy, 37(1), 1-15. https://doi.org/10.1017/ics.2020.21

Yen, Wendy M. 1984. "Effects of Local Item Dependence on the Fit and Equating Performance of the Three-Parameter Logistic Model." *Applied Psychological Measurement* 8 (2): 125–45. https://doi.org/10.1177/014662168400800201

More information
https://oe.cd/4cv

Funded by
the European Union

This technical support project is funded by the Technical Support Instrument, and implemented in cooperation with the European Commission's Directorate-General for Structural Reform Support (DG REFORM)