EQ-5D utilities for patients with nonalcoholic steatohepatitis (NASH): a comparison of historical utilities to a recent real-world multi-country patient survey

PCR3

Jesse C Fishman,[1] Victoria Higgins,[2] James Piercy,[2] James Pike[2]

[1]Madrigal Pharmaceuticals, Conshohocken, PA, USA; [2]Adelphi Real World, Bollington, UK

Originally presented at ISPOR Europe 2023; 12-15 November 2023; Copenhagen, Denmark.

BACKGROUND

- Non-alcoholic fatty liver disease (NAFLD) is a spectrum of chronic liver diseases ranging from non-alcoholic fatty liver to non-alcoholic steatohepatitis (NASH)
- The overall global prevalence of NAFLD has increased significantly in recent years to an estimated 32.4%[1]. NASH, especially at the more advanced stages, can be associated with comorbidities such as cardiovascular disease, obesity, and type 2 diabetes[2]
- NASH and its associated comorbidities can profoundly affect the day-to-day lives of affected individuals
- Utility values such as the EQ-5D[3] are widely used as a measure of health benefits in the health technology assessments (HTAs) used by regulatory agencies to provide guidance on the use of new and existing medicines, products and treatments
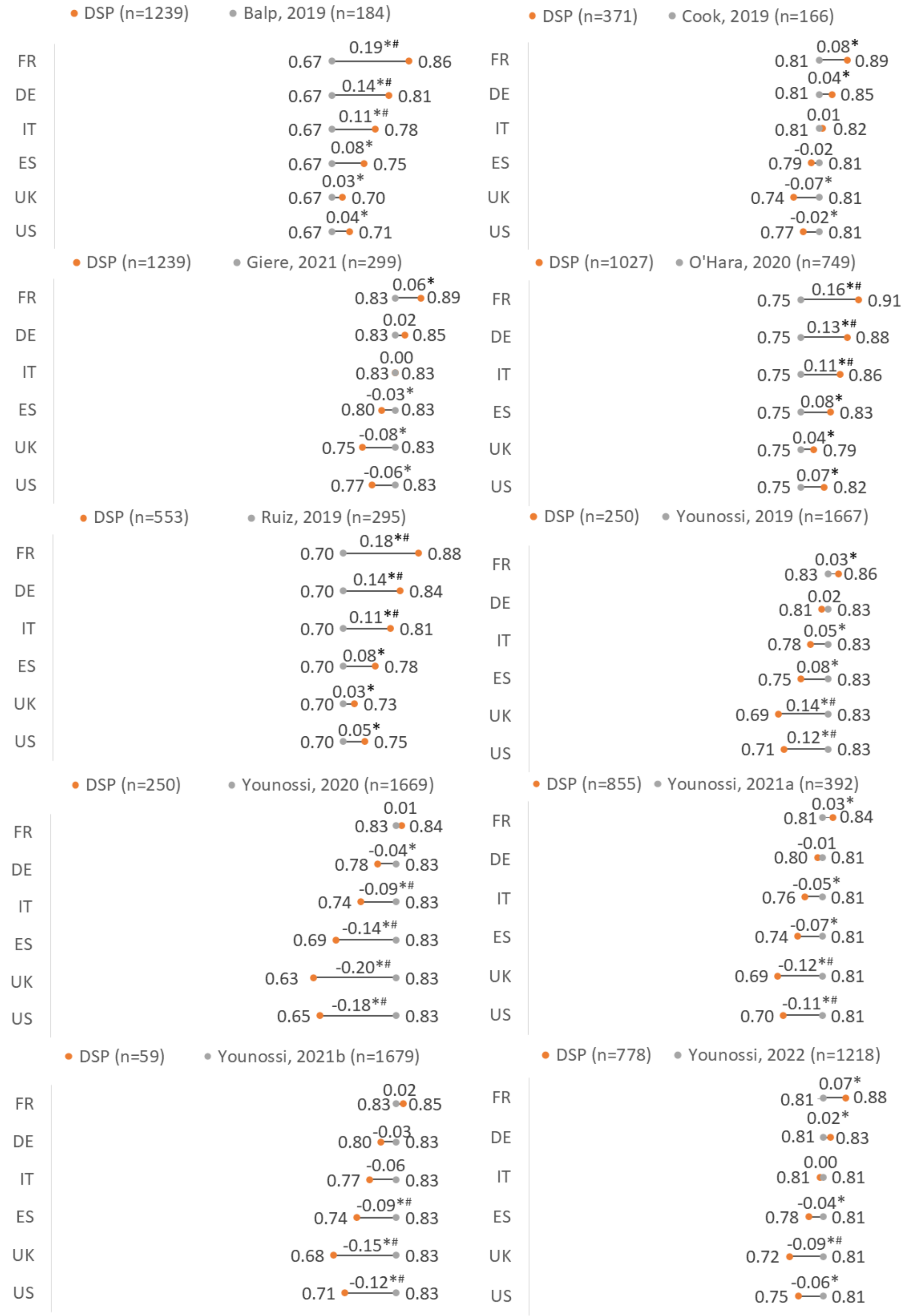
OBJECTIVE

- To describe EQ-5D utilities previously reported in the literature and compare these with real-world EQ-5D utilities for NASH patients from a multi-country patient survey

METHODS

- We performed a structured review for EQ-5D utilities in patients with NASH contained within EMBASE, medline and Cochrane library databases supplemented by searches of key conferences (ISPOR, EASL and AASLD), clinicaltrials.gov and clinicaltrialsregister.eu
- A population, intervention, comparator and outcome methodology[4] was used to find relevant English language studies conducted in any country with no time frame restrictions. Search criteria are shown in Table 1.
- Mean EQ-5D utilities identified in this search were compared with data from the Adelphi Real World NASH Disease Specific Programme (DSP) conducted between January and March 2019 in the United States (US), and between January and March 2018 in France (FR), Germany (DE), Italy (IT), Spain (ES), and the United Kingdom (UK)
- The full methodology of the DSP and its validation has been published [5,6,7]
- For each literature-identified study we:
  - Applied the study's inclusion/exclusion criteria to the DSP data
  - Used matching-adjusted indirect comparison analysis (MAIC) to balance the DSP data with the study
- MAIC is a form of propensity score weighting which can be used to weight patients in a patient-level dataset so that their characteristics exactly match published characteristics
- EQ-5D utilities between weighted DSP and literature values were compared using weighted t-tests , this analysis was repeated for each relevant EQ-5D scoring tariff

Table 1. Systemic literature review search criteria

Search string themes	Data inclusion criteria	Data exclusion criteria
Population	Adult patients with NASH, NAFLD or suspected NASH	
Intervention	Recommended or partially EASL or AASLD recommended treatments (e.g., lifestyle interventions, bariatric surgery, pioglitazone.) Treatments in phase 2+ of development	Treatments not recommended by guidelines, or discontinued from company development pipelines
Comparator	Any	
Outcome	Indications of treatment efficacy, e.g, NASH resolution, liver fat reduction, fibrosis reduction Indications of safety: AE reporting and mortality QoL outcomes	Restrict AE data extraction to five most common AEs reported and % mild, moderate or severe AEs
Study type	Phase 2+ RCTs and predefined subgroup analyses	Systemic literature reviews, phase 1 trials, secondary analyses, post-hoc subgroup analyses, pooled analyses and studies where only 1 arm is of interest

Abbreviations: NASH, Non-alcoholic Steatohepatitis; NAFLD, Non-alcoholic Fatty Liver Disease; QoL, Quality of Life; AE, Adverse event; RCT, Randomised Controlled Trials

RESULTS

- Ten studies with varied recruitment criteria were identified in the systematic literature review with the necessary patient demographic and clinical characteristic data (Table 2)
- In all, 60 t-tests were conducted; with individual comparisons for each of the 10 studies. Analyses was undertaken for each study six times, once for each of the six country tariffs

Table 2. Demographics and Clinical Characteristics

	NASH DSP	Balp et al 2019[8]	Cook et al 2019[9]	Geier et al 2021[10]	O'Hara et al 2020[11]	Ruiz 2019[12]	Younossi 2019[13]	Younossi 2020[14]	Younossi 2021a[15]	Younossi 2021b[16]	Younossi 2022[17]
Country, n (%)											
n	1280	184	17	1216	3754	295	1667	1669	392	1679	1218
France	156 (12.2)	-	-	227 (18.7)	508 (13.5)	-	-	-	-	-	-
Germany	340 (26.6)	-	5 (29.4)	287 (23.6)	540 (14.4)	-	-	-	-	-	-
Italy	120 (9.4)	-	-	-	540 (14.4)	-	-	-	-	-	-
Spain	198 (15.5)	-	-	-	522 (13.9)	-	-	-	-	-	-
UK	102 (8.0)	-	5 (29.4)	-	423 (11.3)	-	-	-	-	-	-
US	364 (28.4)	-	-	702 (57.7)	1221 (32.5)	-	926 (55.5)	1555 (72.2)	336 (85.7)	1306 (60.6)	828 (68.0)
Canada	-	-	6 (35.3)	-	-	-	-	-	-	-	-
Age in years, mean (SD)	55.6 (11.4)	54.5 (13.1)	34–69[a]	54.9 (12.3)	53 (11.9)	-	57.9 (8.8)	-	59.6 (9.0)	57.1 (8.8)	54.1 (11.5)
Sex, n (%)						-					
Male	756 (59.1)	79 (42.9)	8 (47.1)	699 (57.5)	2150 (57.0)	-	673 (40.4)	-	139 (35.5)	854 (39.6)	524 (43.0)
Female	524 (40.9)	105 (57.1)	9 (52.9)	517 (42.5)	1604 (43.0)	-	994 (50.6)	-	-	-	-
Comorbidities, n (%)	1270										
Hypertension	656 (51.7)	91 (49.5)	-	-	-	-	-	-	-	-	-
Heart/blood conditions	678 (53.4)	127 (69.0)	-	-	-	-	-	-	-	-	-
T2DM	729 (57.4)	42 (22.8)	10 (58.8)	-	-	-	1232 (73.9)	1231 (73.8)	282 (71.9)	1555 (72.2)	663 (54.4)
Obesity	720 (56.7)	-	16 (94.1)	676 (55.6)	-	-	-	-	-	-	-
Depression	142 (11.2)	-	5 (29.4)	-	-	-	-	431 (25.8)	-	-	-
Dyslipidemia	589 (46.4)	-	-	-	-	-	-	-	-	-	-
Fibrosis stage, n (%)											
F0	127 (9.9)	-	-	55 (7.0)	F0–2: 35 (11.9)	-	-	-	-	-	287 (23.6)
F1	343 (26.8)	-	-	175 (22.3)	2604 (69.4)	71 (24.1)	-	-	-	-	-
F2	208 (16.2)	-	6 (35.3)	278 (35.4)	-	75 (25.4)	-	-	-	-	411 (33.7)
F3	231 (18.0)	-	5 (29.4)	211 (26.8)	F3–F4: 1150 (30.6)	66 (22.4)	-	-	-	-	520 (42.7)
F4	123 (9.6)	-	-	47 (6.0)	-	48 (16.3)	-	-	-	-	-
Unknown	248 (19.4)	-	-	20 (2.5)	-	-	-	-	-	-	-
Biopsy confirmed NASH, n (%)	653 (51.0)	No	94 (56.6)	786 (64.6)	1619 (43.1)	295 (100)	Yes	Yes	Yes	Yes	Yes
Cirrhosis, n (%)	1052										
Compensated cirrhosis	407 (38.7)	-	-	-	-	-	-	-	-	1133 (52.5)	-
Bridging fibrosis	-	-	-	-	-	-	-	-	-	1021 (47.4)	-
Non-cirrhotic	645 (61.3)	-	-	-	-	-	-	-	-	-	-
Symptoms, n (%)	1269										
Itch/fatigue	-	-	-	-	-	-	-	447 (26.8)	-	-	-
Itch	235 (18.5)	-	-	-	-	-	-	-	-	-	-
Fatigue	832 (65.6)	-	-	-	-	-	-	-	-	-	-
EQ-5D utility score, mean (SD)		0.67 (NR)	0.81 (0.17)	0.83 (0.21)	0.75 (0.26)	0.70 (NR)	0.83 (0.14)	0.83 (0.12)	0.81 (0.14)	0.83 (0.14)	0.81 (0.17)

a Age only reported as a range. Abbreviations: T2DM, Type 2 Diabetes Melitus; NASH, Non-alcoholic Steatohepatitis; SD, Standard Deviation

- Overall, we found that DSP values where significantly higher across all tariffs compared to Balp, 2019, O'Hara, 2020 and Ruiz, 2019, with differences greater than MCID values seen with FR, DE and IT tariffs for all three
- Compared to Cook, 2019 significantly higher utility scores were reported in the DSP when using FR and DE tariffs and significantly lower using US and UK tariffs though none of these were greater than MCID
- Compared to Geier, 2021 DSP reported values where significantly higher when using the FR tariff and lower when using ES, UK and US tariffs though again none of these differences were greater than MCID
- Significant differences were seen across all tariffs except DE compared to Younossi, 2019 and Younossi, 2021a with higher scores reported in the DSP using the FR tariff and lower using IT, ES, UK and US. The UK and US tariffs showed differences greater than MCID
- Meanwhile. compared to Younossi, 2020. all DSP values were significantly lower except when using the FR tariff. IT, ES, UK and US tariffs all produced differences greater than MCID
- Using ES, UK and US tariffs produced DSP scores significantly lower than Younossi, 2021b with differences exceeding MCID in all three
- Compared to Younossi 2022, DSP values were lower with all country specific tariffs applied except IT. Differences were both statistically significant and exceeding the MCID using the UK tariff

LIMITATIONS

- The studies included in this literature review had different inclusion and exclusion criteria, with the result that patients had varying disease characteristics, impacting on their EQ-5D scores. To counter this a range of studies were used to capture a representative breadth of characteristics
- The DSP recruits consecutively consulting patients to reduce physician selection bias; however, this means that those patients who consult most regularly are most likely to be sampled
- Despite these limitations the broad, multi-country, real-world sample allowed for a robust comparison of quality of life in a diverse real-world population

Figure 1. Difference between EQ-5D utility scores: Disease Specific Programme vs Literature

*p<0.05. # difference greater than MCID 0.08
Abbreviations: DSP, Disease Specific Programme' FR, French tariff; DE, German tariff; IT, Italian tariff; ES, Spanish tariff; UK, United Kingdom tariff; United States tariff

CONCLUSIONS

- EQ-5D utilities have been widely reported in the literature but this study has shown that these vary considerably depending on the study populations and methodology
- Assessing differences between values through statistical significance on its own may not be sufficient. Sample sizes can be so large that significant difference are easy to achieve with small differences
- Such small differences may be statistically significant but not clinically meaningful. Therefore, it is also important to consider clinically meaningful differences
- The differences in values observed may be due to inclusion/exclusion criteria, sampling methodology, and the country-specific EQ-5D tariff that was used in the calculation
- Using different country tariffs can produce very different results in patients with NAFLD. This highlights the particular need for country-specific EQ-5D tariffs and sensitivity analysis to confirm the robustness of results and conclusions

References: 1. Riazi K et al (2022) Lancet Gastroenterol Hepatol 7(9): 851-861. 2. Shelley K et al (2023) BMC Gastroenterology 23(1): 160. 3. The EuroQol Group (1990) Health Policy 16(3):199-208 4. Eriksen MB et al. (2018) J Med Libr Assoc 106(4):420-431 5.Anderson P et al (2008) Curr Med Res Opin 24(11): 3063-3072. 6. Babineaux SM et al (2016) BMJ Open 6(8): e010352. 7. Higgins V et al (2016) Diabetes Metab Syndr Obes 9: 371-380. 8.Balp M-M et al (2019) JHEP Reports 1(3):154-161. 9. Cook N et al (2019) Front Med (Lausanne) 6:61 10. Geier A et al (2021) Clin Gastroenterol Hepatol 19(5):1020-1029.e1027. 11. O'Hara J et al (2020) JHEP Reports 2(5):100142. 12. Ruiz L et al (2019) Journal of Hepl (2021) Hepatol Commun 5(7):1201-1211 13. Younossi ZM et al (2019) Clinical Gastroenterology and Hepatology 17(12):2552-2560.e2510 14. Younossi ZM et al (2020) Hepatol Commun 4(11):1637-1650 15. Younossi ZM et al (2021) Hepatol Commun 5(7):1201-1211. 16. Younossi ZM et al (2021) Gastroenterology 160(5):1608-1619.e1613. 17. Younossi ZM et al (2022) Clin Gastroenterol Hepatol 20(9):2050-2058.e2012.