Article published online: 2021-01-19

HyperEmesis Level Prediction (HELP Score) Identifies Patients with Indicators of Severe Disease: a Validation Study

Neuartiges HELP-Instrument zur Identifizierung von Patientinnen mit schwerer Hyperemesis gravidarum: eine Validierungsstudie

OPEN ACCESS

Authors

Kimber W. MacGibbon[1], Sarah Kim[2], Patrick M. Mullin[3], Marlena S. Fejzo[3]

Affiliations

1 HER Foundation, Damascus, OR, USA
2 Brown University, Providence, RI, USA
3 Keck School of Medicine, University of Southern California, Department of Maternal-Fetal Medicine, Los Angeles, CA, USA

Key words

hyperemesis gravidarum, uterus, reproductive medicine

Schlüsselwörter

Hyperemesis gravidarum, Uterus, Reproduktionsmedizin

received 15.6.2020
accepted after revision 10.11.2020

Bibliography

Geburtsh Frauenheilk 2021; 81: 90–98
DOI 10.1055/a-1309-1997
ISSN 0016-5751

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence

Marlena S. Fejzo, PhD
Keck School of Medicine, University of Southern California,
Department of Maternal-Fetal Medicine
2020 Zonal Avenue, #220, Los Angeles, CA 90033, USA
fejzo@usc.edu

Supplementary material is available under https://doi.org/10.1055/a-1309-1997

ABSTRACT

Objective Hyperemesis gravidarum (HG) severity can be underestimated resulting in undertreatment and adverse outcomes. This study was conducted to validate a tool (HELP Score) designed to score HG severity.

Materials and Methods A survey link which included PUQE and HELP Score (HELP) tool questions was posted on websites related to HG. HELP scores were compared to PUQE scores for indicators of severe disease.

Results HELP classified 92% of women reporting "nothing goes or stays down" as severe, compared to 58% using PUQE. Women self-categorizing symptoms as severe were more likely categorized as severe using HELP. Women hospitalized for HG were more likely classified as severe using HELP. HELP performs better than PUQE in identifying patients with severe symptoms requiring intervention.

Conclusion This study provides a novel tool that should be implemented to determine the need for intervention for NVP that may be overlooked using PUQE or empirical assessment.

ZUSAMMENFASSUNG

Zielsetzung Der Schweregrad einer Hyperemesis gravidarum (HG) wird oftmals unterschätzt, was zu einer Unterbehandlung mit negativen Auswirkungen führen kann. Diese Studie wurde zur Validierung eines Instruments (genannt HELP Score [HELP]) durchgeführt, das zur Bewertung des Schweregrads einer HG entwickelt wurde.

Material und Methoden Es wurde ein Link zu einer Publikumsbefragung auf verschiedenen Internetseiten zu HG platziert. Die Befragung bestand aus PUQE- und HELP-Fragen. Die mithilfe des HELP-Instruments erzielten Punktzahlen, die als Indikatoren für eine schwere Erkrankung dienen, wurden mit den erreichten Punkten auf der PUQE-Skala verglichen.

Ergebnisse HELP hat 92% der Frauen, die berichteten, dass „nichts runter geht bzw. nichts unten bleibt", als schwer erkrankt klassifiziert, wohingegen PUQE diese Einstufung nur bei 58% der Frauen vornahm. Mehr Frauen, die ihre Symptome selber als schwer einstuften, wurden mit dem HELP-Instrument als schwer erkrankt eingestuft. Frauen, die wegen HG ins Krankenhaus eingeliefert werden mussten, wurden

eher mit HELP als schwer erkrankt eingestuft. Bei der Identifizierung von Patientinnen mit schweren Symptomen, die ein Eingreifen erfordern, schnitt HELP besser als die PUQE-Skala ab.

Schlussfolgerung Diese Studie beschreibt ein neuartiges Instrument, das zur Ermittlung der Notwendigkeit eines Eingriffs bei schwangerschaftsbedingter Übelkeit und Erbrechen eingesetzt werden kann. Dieses Instrument erfasst auch jene Indikatoren für eine schwere Erkrankung, die beim Einsatz der PUQE-Skala bzw. bei einer empirischen Bewertung oftmals übersehen werden.

Introduction

Nausea and vomiting of pregnancy (NVP) effects 70–90% of all pregnant women, and in the United States as many as 20% of women take medication to treat it. Hyperemesis gravidarum (HG) is at the most severe end of the clinical spectrum, affecting an estimated 2% of pregnant women. HG is the leading cause of hospitalization in the first trimester and is associated with poor maternal and offspring outcomes [1]. Adverse maternal outcomes include severe nutritional deficiencies that can lead to Wernicke's encephalopathy, severe electrolyte imbalance that can result in ventricular fibrillation and cardiac arrest, and prolonged stress that can result in PTSD (DSM-IV-R) following pregnancy. For offspring exposed in utero to HG, there is an increased risk of preterm birth, small for gestational age, low birthweight, and necessity for neonatal care and/or resuscitation [1]. Children are at an increased risk of neurodevelopmental delay and autism spectrum disorder and may be at increased risk of reduced insulin sensitivity, higher blood pressure, increased fat mass and increased cortisol levels [1–3]. Disease severity is often underestimated, resulting in increased maternal weight loss, which may contribute to these adverse maternal and offspring outcomes. Emergency department visits and reports of Wernicke's encephalopathy are on the rise [1]. Of particular concern are reports of maternal deaths in this century due to complications secondary to HG. In a report on maternal deaths in the UK/Ireland, one patient hospitalized "with hyperemesis gravidarum did not see a senior doctor for her entire 5 day admission, and worse, did not see any doctor for the three days prior to her death" [4]. In another maternal death and a recent non-fatal respiratory arrest secondary to HG, miscommunication due to language differences potentially contributed to a delayed understanding of the severity of disease, until it was too late [5]. In addition, women who terminate pregnancies due to debilitating HG were 3 times more likely to report that providers were uncaring or did not understand how sick they were [6]. Suboptimal management is a common concern for HG patients [7]. Therefore, a standardized tool to accurately assess severity of HG is needed to address these potentially preventable outcomes.

The most widely used scoring tool for NVP is the PUQE. The PUQE asks 3 questions pertaining to nausea, vomiting, and retching and is scored from 1–15 with ≥ 13 categorized as severe NVP [8]. Because the PUQE score does not include questions related to quality of life (QOL), a supplemental tool may be required to determine the patient's QOL [9]. The PUQE score may be inadequate for estimating the severity of HG, not only because it does not include QOL, but also because the score is limited to assessing only nausea, vomiting, and retching. Other symptoms contributing to HG and requiring treatment may be overlooked such as poor intake. An example supporting the underestimation of disease severity using the PUQE score comes from a study of women hospitalized with HG in Nepal [10]. In the Nepalese study, women hospitalized for HG for ~ 3.2 days had average PUQE scores at hospital admission that were classified as moderate NVP.

To overcome these issues, the HELP Score tool was created to better assess NVP falling at the severe end of the clinical spectrum, HG. The HELP tool contains questions regarding level of nausea, vomiting, and retching, and adds 9 additional questions valuable in assessing severe disease and necessity for treatment. Included are questions on coping and debility and other questions related to hydration, oral intake, treatment, and progress (▶ **Fig. 1**). Of note, the HELP Score modifies the nausea question in the PUQE because hours of nausea is irrelevant for women with HG – nearly all women with HG have nausea for longer than the maximum on the PUQE (≥ 6 hours). Therefore, nausea severity is assessed instead.

While the PUQE score has been validated to quantify NVP, the use of only 3 questions has the limitation of misclassifying some women with symptoms at the severe end of the clinical spectrum and women requiring immediate intervention. For example, the subset of women who are too nauseous to eat or drink and are bedridden may have low scores for vomiting and retching and thus have low scores on the PUQE, but likely require more aggressive treatment. Therefore, questions relating to intake, keeping food down, urine output, and weight loss can be relevant to severity and were added to the HELP. In addition, questions regarding number of medications, medication tolerance, and symptom improvement after taking medications were added to assess whether patients are able to tolerate and benefit from their treatment regimen. Finally, a question relating to symptom improvement was added to identify patients who may be rapidly deteriorating. The HELP Score ranges from 0–60 with a score of 33–60 classifying patients as severe, providing more room to evaluate levels of severity and improvement in this larger range compared to the 15-point score with only 3-points in the severe range for the PUQE. The HELP scoring tool was subsequently tested using the HG Care Application and was rated to be accurate in defining HG severity by the majority of participating clinicians (100%) and patients (92%) [11]. Herein we compare the PUQE and HELP scores in women diagnosed with HG.

Materials and Methods

HELP Score

Primary clinical symptoms relevant to HG that should be monitored routinely as they indicate or influence severity of the condi-

My nausea level most of the time:	0	1 (mild)	2	3 (moderate)	4	5 (severe)
I average ..... vomiting episodes/day:	0	1–2	3–5	6–8	9–12	13 or more
I retch/dry heave ..... episodes daily:	0	1–2	3–5	6–8	9–12	13 or more
I am urinating/voiding:	Same	More often due to IV fluids; or light color	Slightly less often, and normal color	Once every 8 hours; or slightly dark yellow	Less than every 8 hours or darker	Rarely; dark or bloody; or foul smell
Nausea/vomiting severity 1 hour *after* meds **or** after food/drink if no meds:	0 or no meds	1 (mild)	2	3 (moderate)	4	5 (severe)
Average number of hours I'm *unable* to work adequately at my job and/or at home due to being sick has been:	0	1–2 (hours are slightly less)	3–4 (can work part time)	5–7 (can only do a little work)	8–10 (can't care for family)	11+ (can't care for myself)
I have been coping with the nausea, vomiting and retching:	Normal	Tired but mood is okay	Slightly less than normal	It's tolerable but difficult	Struggling: moody, emotional	Poorly: irritable, depressed
Total amount I have been able to eat/drink **and** keep it down: Medium water bottle/large cup – 2 cups/500 mL	Same; no weight loss	Total of about 3 meals and 6+ cups fluid	Total of about 2 meals and some fluid	1 meal and few cups fluid; or only fluid or only food	Very little, < 1 meal/ minimal fluids; or frequent IV	Nothing goes or stays down, or daily **IV/TPN/NG**
My anti-nausea/vomiting meds stay down or are tolerated:	No meds	Always	Nearly always	Sometimes	Rarely	Never/**IV/SQ** (SubQ pump)
My symptoms compared to last week:	Great	Better	About same	Worse	Much worse	So much worse!!!
Weight loss over last 7 days:	0%	1%	2%	3%	4%	5%
Number of Rx's for nausea/vomiting*:	0	1	2	3	4	5+
	0 pts	**1 pt/answer**	**2 pts/answer**	**3 pts/answer**	**4 pts/answer**	**5 pts/answer**
Total each column = (No. of answers in column) × (No. of points for each answer)	0	.....	.....	.....	.....	.....
Total for all columns:		None/mild ≤ 19	Moderate 20–32		Severe 33–60	

Weight loss (%) = $\left(\frac{\text{amount lost}}{\text{pre-pregnancy weight}}\right) \times 100$

(Weight loss calculation optional for home use)

* Number of Rx's = number of Rx medications for HG (not doses)

IV = Intravenous
TPN = Total parenteral nutrition (IV nutrition)
NG = Nasogastric tube feeding
SQ/SubQ = Subcutaneous
Rx = Prescription medication

▶ **Fig. 1** HyperEmesis Level Prediction (HELP) Scoring Tool.

tion were included in the HELP Score. These included HG symptoms (nausea, vomiting and retching), intake, psychosocial functioning, hydration, treatment effectiveness, and overall progress. Appropriate questions were devised from those clinical indicators in patient-friendly language to assess the current status of the patient. Each question was given 6 possible answers, with the first being a normal or non-HG response and the last one suggesting a very severe condition. Each answer increases in severity. The answers are assigned numerical values from zero to 5, zero for normal and 5 for severe. Answer values are totaled for all questions to get their HELP Score.

Additional questions were added that indicate a higher level of intervention or severity to ensure patients debilitated by the conditions received an appropriate score. These two questions were percent weight loss for last 7 days and number of prescription medications for HG. Women still losing weight or on a large number of medications for nausea and vomiting sometimes will have their primary symptoms reduced, but still be debilitated and unable to meet daily requirements for intake. These reflect a more severe or refractory case, thus including these two questions ensure these patients score appropriately. Each percent weight loss is given one point, with the most severe answer being ≥ 5% and receiving 5 points. Similarly, the number of prescription medications is given 1 point for each medication being used to treat HG, including IV fluids. Those with 5 or more medications receive 5 points.

Survey

A survey was created to identify patients with HG and assess their severity using the PUQE and HELP scoring tools, as well as collect

▶ **Table 1** Demographic characteristics. Demographic characteristics were totaled for each sub-category and % of each category was calculated based on the number included in each sub-category, divided by the total participants (n = 445) with participants with missing data included as a sub-category.

	n	%
Race/ethnicity		
▪ White/Caucasian	347	78.0
▪ Hispanic/Latino	21	4.7
▪ Asian	14	3.1
▪ Black/African American	14	3.1
▪ Mixed	13	2.9
▪ Other	8	1.8
▪ Missing	28	6.3
Education		
▪ Bachelor degree	149	33.5
▪ High school	97	21.8
▪ Graduate degree	97	21.8
▪ Professional certificate	54	12.1
▪ Other	28	6.3
▪ Missing	20	4.5
Age		
▪ 18–23	87	19.6
▪ 24–28	146	32.8
▪ 29–33	130	29.2
▪ 34–38	55	12.4
▪ 39–43	7	1.6
▪ Missing	20	4.5
Medical insurance/pay		
▪ Private insurance	180	40.4
▪ Government/state	159	35.7
▪ Self-pay	18	4.0
▪ Combination of above	50	11.2
▪ Other	36	8.1
▪ Missing	2	0.4
Current work status		
▪ Employed/student full-time	134	30.1
▪ Employed/student part-time	29	6.5
▪ Not employed outside home	82	18.4
▪ Disability/leave due to HG	104	23.4
▪ Left workforce due to HG	55	12.4
▪ Other	21	4.7
▪ Missing	20	4.5

▶ **Table 2** Medications/Treatments. Medications/treatments were totaled for each medication/treatment and % of each category was calculated based on the total number of survey participants that used that medication/treatment, divided by the total participants (n = 445). Participants who did not answer questions regarding their medication/treatment (missing) were included as a sub-category in the total. Number of daily prescribed medications was based only on the 445 participants who answered the question of total number treatments for NVP.

	n	%
Medication/treatment		
▪ Ondansetron (Zofran)	314	71
▪ Acid reducer	129	29
▪ Unisom, Benadryl, Cyclizine, Meclizine	126	28
▪ Promethazine (Phenergan)	97	22
▪ Metoclopramide (Reglan)	86	19
▪ Diclectin/Diclegis	63	14
▪ IV fluids	63	14
▪ Acid blocker	32	7
▪ Marijuna/Marinol	27	6
▪ Compazine/prochlorperazine, Stemetil	26	6
▪ Steroids/methylprednisone	16	4
▪ PICC line/central line	16	4
▪ Home IV therapy	13	3
▪ IV nutrition/TPN	11	2
▪ Gabapentin/neurontin	4	1
▪ NG/NJ feedings	3	1
▪ Kytril/granisetron/Sancuso	1	0
▪ Missing	5	1
Number of daily prescribed medications*		
▪ 0	57	13
▪ 1	102	23
▪ 2	110	25
▪ 3	91	20.4
▪ 4	54	12.1
▪ ≥ 5	31	7

* among those that answered (n = 445)

other demographic characteristics and information on their disease (Supplement 1). The web-based software, Survey Gizmo (Professional version), was used for the survey. It captured the data and allowed setting of appropriate numerical ranges for responses and mandated entry of answers to required questions in order to proceed to the next set of questions. Originally, only the HELP and PUQE scoring questions and a few key health indicators were mandatory, but early on in the process, additional high priority questions were converted to the mandatory setting which is reflected by a system-assigned red asterisk (Supplement 1) to ensure the majority of participants would answer those questions.

Study subjects

Patients were included in the study if they reported HG defined as nausea and vomiting during pregnancy that causes debility and weight loss typically over 5% and requires medical care such as IV fluids and/or prescription medications. The link to the survey was posted on the HER Foundation social media websites between September 6, 2016 and April 23, 2019. Analysis was limited

▶ **Table 3** Medications/treatments by country. For medications/treatments by country, only countries with more than 10 participants were included. Percentages of each medication/treatment per country were based on the total (n) number of participants who answered use of the medication/treatment from that particular country divided by the total number who answered the country of origin question for the country listed. The number of participants who did not answer the medication section of the survey were included in the total as a sub-category and reported for each country.

Country	United States	United Kingdom	Australia	Canada	New Zealand	Nether-lands
n	**233**	**63**	**57**	**18**	**14**	**13**
Ondansetron/Zofran (%)	69	76	77	67	71	46
Reglan/metoclopramide (%)	20	17	19	11	21	0
Kytril/granisetron/Sancuso (%)	0	0	0	0	0	0
Diclectin/Diclegis (%)	16	10	11	11	14	23
Unisom, Benadryl, Cyclizine, Meclizine (%)	23	37	32	50	36	31
Phenergan/promethazine (%)	25	29	14	6	21	0
Compazine/prochlorperazine, Stemetil (%)	5	5	7	0	0	0
Steroids/methylprednisone (%)	3	2	12	0	7	0
Acid reducer (%)	24	30	47	17	21	31
Acid blocker (%)	6	6	7	0	14	8
Marijuana or Marinol (%)	6	2	9	11	0	8
Gabapentin/neurontin (%)	0	2	4	0	0	0
IV fluids (%)	18	13	7	17	21	8
IV nutrition/TPN (%)	2	3	4	0	7	0
Home IV therapy (%)	3	5	0	0	7	8
NG/NJ feedings (%)	0	0	0	0	7	0
PICC line/central line (%)	3	6	0	6	7	0
Missing (n)	3	0	1	1	0	0

to women with HG continuing beyond the first trimester (pregnancy weeks 13–26).

Statistical analysis

Demographic characteristics (race/ethnicity, education, age, medical insurance, and current work status) were totaled for each sub-category and % of each category was calculated based on the number included in each sub-category, divided by the total participants (n = 445) (▶ **Table 1**). Participants with missing data were included as a sub-category. Similarly, medications/treatments were totaled for each medication/treatment and % of each category was calculated based on the total number of survey participants that used that medication/treatment, divided by the total participants (n = 445). Participants who did not answer questions regarding their medication/treatment (missing) were included in the total. The number of daily prescribed medications was based on the 445 participants who answered that question (▶ **Table 2**). For medications/treatments by country (▶ **Table 3**), only countries with more than 10 participants were included. Percentages of each medication/treatment per country were based on the total (n) number of participants who answered use of the medication/treatment divided by the total number who answered the country of origin question for the country listed. The number of participants who did not answer the medication section of the survey were included as a sub-category in the total and reported for each country in ▶ **Table 3**. For pregnancy characteristics shown in ▶ **Table 4**, average and percentages for each characteristic were each calculated based on the number of participants that answered that specific characteristic and the number of participants who skipped that question are shown but not included in the calculation of the percentage for each characteristic.

HELP scores were compared to PUQE scores for participants with answers indicating severe disease requiring intervention including poor urinary output, low food intake, difficulty coping, severe debility, worsening symptoms, inability to take medications, and extreme weight loss (▶ **Table 5**). The number of participants whose HELP and PUQE score categories were discordant was totaled as well as the number of participants whose HELP Score fell in the severe category and PUQE score fell into the moderate or mild category for each severity indicator. To determine whether the HELP Score is better at detecting severe NVP than PUQE, we must reject the null hypothesis that HELP is **not** better at detecting severe NVP than PUQE. Therefore, our null hypothesis is that the PUQE score is at least as good as the HELP Score at detecting severe NVP and data is analyzed using a one-sided sign test. The assumption is made that each participant with an indicator of severe disease has severe NVP. The binomial distribution is used to estimate the maximum probability that, given the null hypothesis is true, we end up with the results shown in ▶ **Table 5**.

Mean PUQE and HELP scores were calculated for those participants who reported having disease indicators that require(d) immediate intervention including very little or no meal intake, uri-

▶ **Table 4** Pregnancy Characteristics. For pregnancy characteristics, average and percentages for each characteristic were calculated based on the number of participants that answered that specific characteristic and the number of participants who skipped that question are shown but not included in the calculation of the percentage for each characteristic.

Characteristic among those that answered		Missing
Average number of HG pregnancies	1.9	76
Reporting on first pregnancy with HG	40%	76
Reporting on recurrent HG pregnancy	56%	76
Average full-term deliveries	1.22	154
Average miscarriages/stillbirths	1.27	233
Therapeutic (due to HG) termination	27%	349
Considered terminating due to HG	41%	0
Average weight loss*	13%	0
Extreme weight loss (≥ 15%)	18%	0
ER visits	59%	5
Inpatient hospitalization	45%	5
No vitamins or supplements in past 24 hours	46%	32

* weight loss from pre-pregnancy weight

nating rarely, rarely or never tolerating nausea medication, and having at least one previous emergency department visit or hospitalization (▶ **Table 6**). In addition, among participants reporting the listed disease indicator, the percentage of participants scoring moderate and percentage scoring severe by PUQE and by HELP for each indicator are shown.

For analysis of self-reported severity of NVP compared to PUQE and HELP Score, the p-value and odds ratio values were calculated using medcalc.org odds ratio calculator to determine the likelihood that the HELP Score agreed with self-reported severity compared to the likelihood that the PUQE score agreed with self-reported severity. This study has been approved by UCLA ORA, IRB#11-001681.

Results

Four-hundred forty-five women diagnosed with HG and in the 2nd trimester of pregnancy completed the survey questions required to generate a HELP and PUQE score and were included in the study.

Demographic characteristics

Demographic characteristics are shown in ▶ **Table 1**. The majority of participants were white, with a bachelor's degree, aged 24–28, with private health insurance, and employed or a full-time student.

Medications and treatments

Medications and treatments are shown in ▶ **Table 2** with the most common medication being ondansetron. Almost 40% of participants reported taking 3 or more prescription medications for NVP daily, with 25% of participants taking 2 prescribed medications, 13% taking no prescription medications, and 7% taking 5 or more prescription medications. Women from 25 countries participated in the survey. Among these, there were six countries with more than 10 participants (United States, United Kingdom, Australia, Canada, New Zealand, and the Netherlands) and these participants were included in the medications/treatments broken down by country, shown in ▶ **Table 3**. Ondansetron was the most common medication in all 6 countries, ranging from 46% in the Netherlands to 77% in Australia. The 2nd most common treatment varied. For example, Promethazine (Phenergan) was the 2nd most common treatment in the United States used to treat 25% of women, while the second most common treatment in Canada, the UK, and New Zealand was antihistamines (Unisom, Benadryl, Cyclizine, Meclizine) used to treat 50, 37, and 36% respectively. Acid reducers were the 2nd most common treatment (47%) in Australia. In the Netherlands 31% of women reported treatment with antihistamines and/or acid reducers. Marijuana/marinol to treat HG ranged from 0% in New Zealand to 11% in Canada. New Zealand had the highest percentage (7%) of participants reporting more aggressive treatments such as home IV therapy and tube feeding.

Pregnancy characteristics

Pregnancy characteristics are shown in ▶ **Table 4** among participants who answered the question pertaining to the specific characteristic listed in the table. There was a high degree of missingness for some characteristics that were not included as mandatory. Participants had, on average, approximately two HG pregnancies with 56% reporting from 2 to as many as 11 HG pregnancies. Participants had on average 1.2 full-term deliveries and 1.3 miscarriages/stillbirths. While 41% of women considered terminating due to HG, 27% (26/96, 349 missing) reported having had a therapeutic termination due to HG. The average weight loss was greater than 10% of pre-pregnancy weight (12.5%), with extreme weight loss (> 15%) reported in 18% of participants. Over half the participants (59%) had at least one emergency room visit and almost half (45%) were hospitalized. Almost half of the women (46%) had not taken any vitamins or supplements in the past 24 hours. And less than 15% of participants reported that they were able to take and keep down prenatal vitamins every day for the last 2 weeks.

Comparison of HELP and PUQE scores for indicators of severe disease

HELP Score indicators of severe disease requiring intervention include patients reporting any of the following: urinary output – rarely, intake less than or equal to 1 meal, struggling/poor coping, significant debility (can't care for self), symptoms worsening, medication tolerance rare/none, having lost a lot of weight in the last week, and extreme weight loss (greater than 15% of pre-pregnancy weight). ▶ **Table 5** shows that women reporting any of these were classified significantly differently using the PUQE score compared to the HELP Score. Women with these symptoms were more likely to be classified as having severe HG requiring more aggressive treatment using the HELP scoring tool (▶ **Tables 5** and **6**). For example, 75% of women who reported rarely any urinary out-

▶ **Table 5** Comparison of HELP vs. PUQE for indicators of severe disease. HELP scores were compared to PUQE scores for participants with answers indicating severe disease requiring intervention. The number of participants whose HELP and PUQE score categories were discordant was totaled as well as the number of participants whose HELP Score fell in the severe category and PUQE score fell into the moderate or mild category for each severity indicator. To determine whether the HELP Score is better at detecting severe NVP than PUQE, we must reject the null hypothesis that HELP is **not** better at detecting severe NVP than PUQE. Therefore, our null hypothesis is that the PUQE score is at least as good as the HELP Score at detecting severe NVP and data is analyzed using a one-sided sign test. The assumption is made that each participant with an indicator of severe disease has severe NVP. The binomial distribution is used to estimate the maximum probability that, given the null hypothesis is true, we end up with the results shown.

Indicator of severe disease	HELP and PUQE disagree (n)	SEVERE with HELP but not PUQE (n)	p-value
Urinary output – rarely	3	3	0
Eat/drink < 1 meal	58	44	5.02E-05
Struggling/coping – poorly	66	51	5.05E-06
Debility – can't take care of self	28	25	1.37E-05
Symptoms worsening	32	26	2.68E-04
Medications rarely/never stay down	22	19	4.28E-04
Lost a lot of weight this week	15	15	0
Extreme weight loss	11	9	3.27E-02

▶ **Table 6** PUQE score is more likely to misclassify women requiring immediate intervention as "moderate NVP." Mean PUQE and HELP scores were calculated for those participants who reported having disease indicators that require(d) immediate intervention. Among participants (n) reporting the listed disease indicator, the percentage of participants scoring moderate and percentage scoring severe by PUQE and by HELP for each indicator are shown.

Disease indicator requiring/required intervention	n	PUQE score mean (≥ 13 = SEVERE)	HELP Score mean (≥ 33 = SEVERE)	% scored MODERATE using PUQE	% scored MODERATE using HELP	% scored SEVERE using PUQE	% scored SEVERE using HELP
Meal intake (eat nothing)	12	12.8	44.3	41.7	8.3	58.3	91.7
Meal intake (very little or nothing)	59	12.6	40.1	44.1	13.6	55.9	86.4
Urinating rarely	8	10.9	37.6	75	37.5	25	62.5
Rarely or never tolerate NVP medication	72	11.9	37.2	52.8	30.6	47.2	69.4
≥ 1 Emergency department visit for HG	265	10.1	28.2	65.7	46.8	23.4	33.6
≥ 1 Hospitalization for HG	202	9.8	27.4	66.8	45	18.8	31.7

put were classified as having moderate NVP using the PUQE score compared to 38% classified as moderate HG using the HELP Score, and 42% of women reporting eating nothing scored moderate NVP on the PUQE compared to 8% scoring moderate HG using the HELP tool.

For women reporting that they rarely or never tolerate/keep down their antiemetic medication, the average PUQE score was 11.9 (moderate NVP), while the average HELP Score was 37.2 (severe HG). Thus, the average woman who is not taking her medication and likely in need of a different medication or form of administration was classified as having moderate NVP using the PUQE but severe HG using the HELP tool. This was also the case for extremely poor intake and rare urinary output (▶ **Table 6**) suggesting clinician decision-making would benefit from usage of the HELP. Using indicators of severe disease that are not included in the HELP scoring tool such as having at least one emergency department visit or at least one hospitalization due to HG, participants were also more likely to score severe using the HELP tool compared to the PUQE tool (▶ **Table 6**). For example, among the 445 participants, 440 answered questions on hospitalization (and emergency department visits), and 18.8% of women who were hospitalized at least once for HG scored severe using the PUQE compared to 31.7% using the HELP scoring tool.

Comparison of HELP and PUQE scores and self-classification of disease severity

Finally, in this study, patients were also asked to categorize their own NVP symptoms (defined in Supplement 1, question 39). All 445 participants answered this question. Ten percent (n = 43) of the participants self-categorized their symptoms as severe (unable to function and constantly very sick) and among those, 70% were categorized as severe using the HELP Score, but only 47% were categorized as severe using the PUQE score. Women who self-categorized themselves as having severe symptoms were 2.7-fold more likely to score severe using the HELP tool than the PUQE tool (OR = 2.7, p = 0.03). On the other end of the spectrum,

there was not a significant difference between HELP and PUQE scores in categorizing women as severe when they self-categorized themselves as mild. In this study, 42% of study participants categorized their NVP symptoms as mild, and among those, 8% were categorized as severe by the HELP Score and 6% were categorized as severe using the PUQE score.

Discussion

This study validates the HELP Score as a better tool than the PUQE at identifying women at the severe end of the clinical spectrum of NVP/HG.

There remains a critical need for better tools to diagnose, manage, and treat HG, despite recent progress suggesting the placenta and appetite hormone GDF15 plays a role in the etiology of the condition [1, 12 – 15]. HG can be associated with extreme weight loss and nutritional deficiencies that may lead to adverse maternal, fetal, and child outcomes. Even in this century among patients under hospital care, underestimation of severity of disease has resulted in maternal deaths in the UK, USA and other countries [5] (MSF, personal communication). Undertreatment has also resulted in pregnancy termination – 26 women in this study reported at least one pregnancy termination due to HG. HG is associated with preterm birth, small for gestational age, and low birth weight, especially when there is prolonged symptoms and inadequate weight gain [1]. The HELP tool is more effective at identifying women in need of intervention and classifying them as severe which may reduce these adverse outcomes. For example, the HELP tool classified the majority (92%) of women who reported "nothing goes or stays down" as severe HG, compared to only 58% using the PUQE. Furthermore, women self-categorizing their symptoms as severe were significantly more likely to be categorized as severe using the HELP Score than the PUQE.

Emergency department visits for HG are on the rise with over 391 000 visits per year in the United States in 2014 [16]. Over half (60%) of the participants in this international study had at least one emergency department visit and 45% were hospitalized at least once for HG during some week of their pregnancy. Women who had visited the emergency room or who had been hospitalized at least once in their pregnancy for HG were also more likely to be classified as severe using the HELP tool compared to the PUQE.

Among the top 6 participating countries, the most commonly prescribed medication (71%) was ondansetron. Almost 40% of women reported taking 3 or more prescription medications for NVP daily. However, a recent study showed almost two-thirds of women who are hospitalized for HG are not prescribed an antiemetic prior to admission [17]. Use of marijuana/marinol to treat symptoms ranged from 0% in New Zealand where more aggressive interventions were reported, to as high as 11% in Canada. Regular assessments using the HELP tool to screen women with NVP symptoms at home, at provider visits, and in the emergency department/hospital setting worldwide may lead to improved intervention and prevent women from being undertreated and/or turning to alternative therapies with less safety data. Of particular concern, 43% of women reported taking no vitamins or supplements in the last 24 hours which can put them at risk for Wernicke's encephalopathy and may have long-term effects on outcome, especially in the subset of women reporting little or no intake.

Importantly, the HELP Score tool has been implemented in the free HG Care iPhone application, which has already been shown to improve patient-provider communication and care in a beta testing study [11]. Future studies should focus on determining whether implementation of the HELP tool effectively reduces number of prescription medications, emergency department visits, hospitalization, readmissions, pregnancy terminations due to HG, preterm births, and other adverse maternal and offspring outcomes, including neurodevelopmental delay and autism spectrum disorder. Implementation of the HELP Score tool worldwide may also solve the problem of heterogeneity in HG definitions limiting meta-analyses of HG studies, which has been a major factor inhibiting progress in HG research [1].

Strengths and Limitations

Admittedly, this study has some weaknesses. For example, the study is based on self-reports and medical records were not collected. Some answers had a high degree of missingness or were changed to be mandatory early on in the study, and therefore cannot be generalized due to selection bias and should be interpreted with caution. For example, 349 women did not answer the question on therapeutic termination due to HG, and therefore while 27% of the population that answered the termination question reported having a termination, this (27%) can only be interpreted as descriptive of the population that answered the question, but is likely a gross overestimate of the actual termination rate. The actual termination rate is likely closer to 6% based on a previous study as well as in this study, when assuming all participants who skipped the question did not have a therapeutic termination [18]. The participants came from social media websites and therefore this patient pool may not be generalizable to the entire population. However, the majority of questions were specifically based on the participants' experiences in "the last 24 hours" and therefore recall is not likely to be an issue. In addition, since all English-speaking participants with internet had access to the survey, it may be more generalizable than a study performed in a single setting. Further support for this, and a strength of the study, is that participants came from 25 countries across a wide age range with multiple ethnicities and educational backgrounds represented. Another limitation is that the analysis was limited to women with symptoms beyond the 1st trimester. We chose to focus on the most clinically vulnerable subgroup with respect to poor maternal and child outcomes-women who are more likely to have prolonged symptoms and inadequate weight gain [1]. Data was collected in the 1st trimester with comparable findings (not shown), and the results are generalizable to women in their first trimester.

Finally, the HELP Score is designed to evaluate women with severe NVP or HG, and its utility in the general population with normal nausea and vomiting of pregnancy has not been addressed in this study. However, women scoring less than 20 on the HELP are determined to have none or mild HG and likely do not require changes to medical intervention such as additional antiemetic medication or intravenous fluids for NVP.

Conclusion

The HELP Score is a valid 12-question tool for screening and scoring levels of severe NVP. It should replace the PUQE in assessing women with NVP that are suspected to require interventions as they may be missed by the PUQE or empirical assessment.

Funding

There is no funding to report.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Fejzo MS, Trovik J, Grooten IJ et al. Nausea and vomiting of pregnancy and hyperemesis gravidarum. Nat Rev Dis Primers 2019; 5: 62

[2] Getahun D, Fassett MJ, Jacobsen SJ et al. Autism Spectrum Disorders in Children Exposed in Utero to Hyperemesis Gravidarum. Am J Perinatol 2019. doi:10.1055/s-0039-1696670

[3] Poeran-Bahadoer S, Jaddoe VWV, Gishti O et al. Maternal vomiting during early pregnancy and cardiovascular risk factors at school age: the Generation R Study. J Dev Orig Health Dis 2020; 11: 118–126

[4] Knight M, Kenyon S, Brocklehurst P, Neilson J, Shakespeare J, Kurinczuk JJ, eds.; on behalf of MBRRACE-UK. Saving Lives, Improving Mothers' Care – Lessons learned to inform future maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2009–12. Oxford: National Perinatal Epidemiology Unit, University of Oxford; 2014

[5] Fejzo MS, MacGibbon K, Mullin PM. Why are women still dying from nausea and vomiting of pregnancy? Gynecol Obstet Case Rep 2016; 2: 2

[6] Poursharif B, Korst LM, Macgibbon KW et al. Elective pregnancy termination in a large cohort of women with hyperemesis gravidarum. Contraception 2007; 76: 451–455

[7] Havnen GC, Truong MB, Do MH et al. Women's perspectives on the management and consequences of hyperemesis gravidarum – a descriptive interview study. Scand J Prim Health Care 2019; 37: 30–40

[8] Koren G, Piwko C, Ahn E et al. Validation studies of the Pregnancy Unique-Quantification of Emesis (PUQE) scores. J Obstet Gynaecol 2005; 25: 241–244

[9] McParlin C, Carrick-Sen D, Steen IN et al. Hyperemesis in Pregnancy Study: a pilot randomised controlled trial of midwife-led outpatient care. Eur J Obstet Gynecol Reprod Biol 2016; 200: 6–10

[10] Chhetry M, Thakur A, Uprety DK et al. Hyperemesis Gravidarum in a Tertiary Care Centre in Eastern Nepal: A Prospective Observational Study. J Ayub Med Coll Abbottabad 2016; 28: 18–21

[11] Korouri E, MacGibbon K, Chan M et al. Performance of iPhone Hyperemesis Gravidarum Care App. J Clinical Case Rep Case Stud 2019; 2: 53–59

[12] Petry CJ, Ong KK, Burling KA et al. Associations of vomiting and antiemetic use in pregnancy with levels of circulating GDF15 early in the second trimester: A nested case-control study. Wellcome Open Res 2018; 3: 123

[13] Fejzo MS, Sazonova OV, Sathirapongsasuti JF et al.; 23andMe Research Team. Placenta and appetite genes GDF15 and IGFBP7 are associated with hyperemesis gravidarum. Nat Commun 2018; 9: 1178

[14] Fejzo MS, Arzy D, Tian R et al. Evidence GDF15 Plays a Role in Familial and Recurrent Hyperemesis Gravidarum. Geburtshilfe Frauenheilkd 2018; 78: 866–870

[15] Fejzo MS, Fasching PA, Schneider MO et al. Analysis of GDF15 and IGFBP7 in Hyperemesis Gravidarum Support Causality. Geburtshilfe Frauenheilkd 2019; 79: 382–388

[16] HCUPnet. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. Accessed September 23, 2020 at: http://www.hcup-us.ahrq.gov/

[17] Fiaschi L, Nelson-Piercy C, Deb S et al. Clinical management of nausea and vomiting in pregnancy and hyperemesis gravidarum across primary and secondary care: a population-based study. BJOG 2019; 126: 1201–1211

[18] Fejzo MS, MacGibbon KW, Mullin PM. Ondansetron in pregnancy and risk of adverse fetal outcomes in the United States. Reprod Toxicol 2016; 62: 87–91