Coloplast

Clinical Investigation Report

CP345 The ASSISTER Trial

A rAndomised croSS-over trIal inveSTigating HEylo™, a novel app-driven digital supporting ostomy pRoduct

January 2022 – November 2022

Investigator

PRINCIPAL INVESTIGATOR	CONTACT INFORMATION
[redacted]	[redacted]

Sponsor

Coloplast A/S
Holtedam 1
DK-3050 Humlebæk
Denmark
+45 4911 1111

1 Table of Contents

1.1 List of tables

1.2 List of figures

2 Summary

Title	Time schedule	ITT population
CP345 The ASSISTER Trial – A randomised cross-over trial investigating Heylo™, a novel app-driven digital supporting ostomy product.	Initiation date (first subject in): 13th of January 2022 Completion date (last subject out) dates: 4th of November 2022	n=139

Introduction

People with intestinal stomas can have problems with leakage which negatively influence their quality-of-life (QoL), despite the development of better ostomy products. In a recent international survey, it was reported that 26% of the people living with a stoma had experienced leakage outside the baseplate monthly or more frequently and 66% had experienced this at least once during the past year.

To overcome the burden of leakage, Coloplast has developed a novel app-driven digital leakage notification system for ostomy care called Heylo™. Heylo™ consists of three parts; a smartphone software application, which is installed on the user's smartphone (The Heylo™ app), a sensor layer and a transmitter. The Heylo™ app informs the user of any changes of the user's baseplate status, based on measurements obtained from the adhesive sensor layer placed underneath the baseplate where it continuously monitors for stomal effluent leakage and moisture. The transmitter, which is connected to the sensor layer, continuously evaluates the information from the sensor layer and sends the status to the Heylo™ app.

Aim and objectives

We hypothesized that the novel digital leakage notification system (Heylo™) is associated with positive care effects in people with intestinal stomas. The aim of this trial was therefore to investigate the effect of Heylo™ on QoL and disease burden in ostomy care.

The **primary objective** was to investigate whether Heylo™ can improve the *Emotional impact* domain score of the Ostomy Leakage Impact tool (OLI), compared to standard of care after 8 weeks of product usage.

The **secondary objective** was to evaluate whether Heylo™ can improve participation in everyday- and social activities measured by the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) *Participation* domain (Domain 6) compared to standard of care after 8 weeks of product usage.

Endpoints and methods

Primary endpoint: *Emotional impact* domain score (scale from 0–100, the higher scores corresponding with better emotional impact) measured by the OLI tool evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).

Secondary endpoint: *Participation* domain score (scale from 0 – 100, the lower scores corresponding with better participation in everyday- and social activities) measured by the WHODAS 2.0 evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).

The estimated treatment effect, based on either the primary and/or secondary endpoint, should be sufficient on its own (tested on equal terms) to establish the positive care effect. Hence, a Bonferroni correction to account for the familywise type 1 error, for two endpoints was applied. The corresponding acceptance level for either endpoint should be at a significance level of $p<0.025$ (standard significance level of $p<0.05$ corrected for multiple testing of two endpoints).

Exploratory endpoints:

- *Usual and social activities* domain score (scale from 0-100) measured by the OLI scale evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- *Coping and in Control* domain score (scale from 0-100) measured by the OLI scale evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- Cognition domain score (scale from 0-100) measured by WHODAS 2.0 evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- Mobility domain score (scale from 0-100) measured by WHODAS 2.0 evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- *Self-care* domain score (scale from 0-100) measured by WHODAS 2.0 evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- *Getting along* domain score (scale from 0-100) measured by WHODAS 2.0 evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- *Life activities* (household and work/school) domain score (scale from 0-100) measured by WHODAS 2.0 evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- Feeling of security evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).

Additional exploratory endpoints (described in the statistical analysis plan)

- Total WHODAS 2.0 score (mean of all 6 domain scores) (scale 0-100) evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- Days affected by disability, measured by three additional WHODAS 2.0 questions (Questions H1–H3) evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).

Other assessments and safety assessments:

- Leakage outside the baseplate, evaluated at the end of each test period (after 8 weeks on Heylo or standard of care).
- Change in current ostomy solution, evaluated at the end of each test period.
- Change in Heylo™ size, evaluated at the end of each test period.
- Adverse events (AEs)/device deficiencies

Study design

This was an open-label, randomized controlled cross-over trial comparing Heylo™ and standard of care.

Study visits were conducted at the subject´s home or through remote virtual calls.

Each of the subjects had an inclusion visit, a baseline visit (V0 and V1) and two test visits - V2 (after 8 weeks, period 1) and V3 (after 8 weeks, period 2). V3 was the termination visit unless a situation occurred where a subject terminated earlier than expected. If this was the case the subject had, as the last visit, the termination visit performed. All visits were carried out by the Principal Investigator, or delegate.

At V1, V2 and V3 the Principal Investigator or delegate instructed the subjects in completing questionnaires including the WHODAS 2.0 and OLI tools and every 2nd week the subject had to complete the OLI tool and questions about leakage episodes outside the baseplate. All questionnaires were completed using an electronic data-capturing system.

A follow-up call was scheduled 7 days ± 2 days after visit 1 and visit 2 to ensure compliance with the provided investigational device, study procedures and assurance of the subject's wellbeing. Additional calls/visits were scheduled if needed, assessed by the principal investigator or delegates, and registered as unscheduled visits.

Results

The statistical results were based on data from 139 subjects in the intention-to-treat (ITT) population (full analysis set) and the safety summary is based on data from the 144 subjects in the Safety population. The results described below are the primary comparisons performed after 8 weeks use of Heylo™ compared to 8 weeks with standard of care.

Primary endpoint
A significant improvement in the mean *Emotional impact* domain score was found with Heylo™ compared to standard of care after 8 weeks of treatment ($p<0.001$).

Secondary endpoint
A significant improvement in the mean *Participation* domain score was found with Heylo™ compared to standard of care after 8 weeks of treatment ($p=0.001$).

Explorative endpoints
The mean score for the two additional OLI domains *Usual and social activities* and *Coping and in Control* also improved significantly with Heylo™ compared to standard of care after 8 weeks of treatment ($p<0.001$, respectively).

Except for the *self-care* domain score, statistically significant improvements were seen in all WHODAS 2.0 domains and the mean of the total WHODAS 2.0 score with Heylo™ after 8 weeks of treatment, when compared to standard of care (D1: $p=0.018$, D2: $p=0.014$, D3: $p=0.120$, D4: $p=0.020$, D5: $p=0.008$, D6: $p=0.001$, Total WHODAS 2.0 $p<0.001$).

For the three additional questions (H1-H3) assessing the days affected by disability the results were as followed: H1: Days with difficulties present (out of 30 days) after 8 weeks of treatment were significantly reduced with Heylo™, compared to standard of care ($p=0.025$). No significant differences were found in H2: Days of being totally unable to carry out activities or work because of any health condition (out of 30 days) ($p=0.543$) and H3: Days with cut back or reduction of usual activities or work (out of 30 days) ($p=0.248$).

The feeling of security increased significantly after 8 weeks of treatment when using Heylo™ compared to standard of care ($p<0.001$). Lastly, a significant reduction in episodes of leakage outside the baseplate was observed with Heylo™ when compared to standard of care after 8 weeks of treatment ($p<0.001$).

Post-hoc analyses
Post-hoc analyses of the primary endpoint and the secondary endpoint for all randomized subjects based on multiple imputation were performed. The results supported a statistically significant treatment effect of Heylo versus standard of care after 8 weeks of treatment from the pre-specified analyses.

Safety assessments:
In all, 5 AEs in 5 different subjects (5/144 *100 = 3.5%) were related to the investigational device, however, none of these were classified as serious AEs.

All the AEs (n=5) related to the investigational device were associated with skin and subcutaneous tissue disorders (primarily skin irritation). For four AEs that were related to the investigational device, the intensity was considered moderate, whereas one AE was considered severe (contact dermatitis).

Hence, it can be concluded that the current investigational device showed no unanticipated AEs considering that the subjects tried out a new type of product with different adhesive area and material.

Conclusion

This randomised controlled cross-over trial demonstrated that Heylo™ provided positive care effects to QoL and to the overall burden of living with a stoma after 8 weeks of product usage, reflected as less embarrassment, better sleep, living with dignity, and better capability of participating in society and interacting with close family and friends.

More specifically Heylo™ demonstrated significant improvements in all three OLI domain scores and in 5 out of 6 WHODAS 2.0 disability domain scores, together with an overall improvement in the total WHODAS 2.0 score and a reduced number of days with difficulties present. Also feeling of security and episodes of leakages outside the baseplate improved significantly with Heylo™ compared to standard of care.

Together, these findings suggest that Heylo™ provides clinically relevant and meaningful positive care effects for people living with a stoma. This intriguing finding was seen in study participants with both ileostomy and colostomy.

3 Abbreviated terms and definitions

ABBREVIATION	WRITTEN OUT	EXPLANATION
ADE	Adverse Device Effect	
AE	Adverse Event	
App	Mobile application	
ASADE	Anticipated Serious Adverse Device Effect	
CI	Confidence interval	
CIP	Clinical Investigation Plan	
CRF	Case Report Form (paper or electronic)	Questionnaire to be used for data collection
CM	Clinical Manager	
DD	Device deficiency	
EC	Ethics Committee	
IFU	Instruction For Use	
ITT	Intention-to-treat	Defined as the full analysis set constituted of all randomised subjects with valid informed consent who have been exposed to at least one product, with information on at least one endpoint
LS mean	Least square mean	Mean estimated from a statistical model by a least squared method
MCID	Minimal clinically important difference	
OLI	Ostomy Leak Impact	Tool to assess the burden of leakage
PI	Principal Investigator	Qualified person responsible for conducting the clinical investigation at an investigation site. If the clinical investigation is conducted by a team of individuals at an investigation site, the PI is the responsible leader of the team. Whether this is the responsibility of an individual or an institution can depend on national regulations.
PP	Per Protocol	
QoL	Quality of Life	
SADE	Serious Adverse Device Effect	
SAE	Serious Adverse Event	
SAP	Statistical Analysis Plan	
SoC	Standard of Care	
USADE	Unanticipated Serious Adverse Device Effect	
WHODAS 2.0	The World Health Organization Disability Assessment Schedule 2.0	Tool to assess health and disability across cultures

4 Introduction

Despite the continuous development of new ostomy product solutions, many people with intestinal stomas still struggle with episodes of leakage progressing outside the baseplate soiling clothes or bed sheets. In a recent international survey, it was reported that 26% of the people living with a stoma had experienced leakage outside the baseplate monthly or more frequently, and 66% had experienced this at least once during the past year (1).

Leakage of stomal effluent progressing outside the baseplate can be socially embarrassing, is often distressing and is always inconvenient for the individual to experience, with more than 90% of people living with a stoma worrying about leakage (2, 3). Multiple studies have highlighted that the frequency of which subjects experience leakage is associated with reduction in quality-of-life (QoL) (3-5) and disutility (6). Subjects use different means to mitigate the risk of experiencing future leakage incidents, such as increasing the use of ostomy product solutions (bags, baseplates and accessories), and some have consultations with health professionals (7).

To overcome the burden of leakage, Coloplast has developed a novel app-driven digital leakage notification system for ostomy care called Heylo™. Heylo™ consists of three parts: a smartphone software application which is installed on the user's smartphone (The Heylo™ app), a sensor layer and a transmitter.

The Heylo™ app informs the user of any changes to the user's baseplate status, based on measurements obtained from the adhesive sensor layer placed underneath the baseplate, where it continuously monitors for stomal effluent leakage and moisture. The transmitter, which is connected to the sensor layer, continuously evaluates the information from the sensor layer and sends the status to the Heylo™ app.

Three previous investigations have already been conducted to refine Heylo™ (CP278, CP308 and CP321) (8-10). The first investigation (CP278) did not test the actual Heylo™ device, but only the sensor technology. It was observed that leakage underneath the baseplate could be detected with high accuracy, yet the investigation revealed that the subjects experienced too many app notifications (8). The sensor system and algorithm for notifications were further developed based on these findings. The modified algorithm and the Heylo™ sensor system were then evaluated in the investigation CP308, which showed that the sensor system performance was acceptable and messages to the subjects' phones were overall evaluated as reliable (9). Still, optimization of the system, including improvements of the Bluetooth and sensor performance was necessary. Firmware and software have since been updated accordingly and the sensor performance algorithm has been optimized based on data from the CP308 investigation (9). In addition, the shape of the sensor layer ink-print has been changed to improve the robustness of the system's leakage detection. The exploratory clinical investigation CP321 was conducted to evaluate the updated system (10). This trial included 25 subjects in Denmark and confirmed technical readiness with both Android and iOS software. The investigation also showed that the use of Heylo™ significantly decreased the number of leakage incidents progressing outside the baseplate and provided clinical benefits to the users, e.g. reduction in users' worry about leakage and improvements in their QoL (10).

Subsequently, two investigations were planned; one with Heylo™ delivered with a support service and one investigating Heylo™ without a support service. The first investigation was a single-arm trial (CP340), planned and conducted in the United Kingdom (UK). The results of this investigation highlighted multiple clinical benefits of Heylo™ when delivered together with a support service, such as reduced incidents of leakages outside baseplate onto clothes, improvements in QoL, as well as increased knowledge, skills, and confidence in managing own health (11).

In this second randomized controlled cross-over trial (CP345) we hypothesized that the novel digital leakage notification system (Heylo™ delivered without a support service) is associated with positive care effects in people with intestinal stomas. The aim of the trial was therefore to investigate the effect of Heylo™ on QoL and disease burden in ostomy care.

The investigation was conducted in Germany in a homecare setting and the target population was people with either an ileostomy or colostomy who had liquid and/or mushy output and who had leakage worries and leakage problems.

5 The investigational device, standard of care and methods

5.1 The investigational device description

The investigational device is the novel app-driven digital leakage notification system called Heylo™.

Heylo™ obtained the European market clearance (CE-mark) in June 2021 (12).

The Heylo™ investigational device consists of the following (Figure 1 and Figure 2):

- The Heylo™ app (installed on the user´s own smartphones) (Figure 1). The app is communicating the sensor status to the user.
- Adhesive patch-sensor layer (single use)
- Transmitter (re-use), to be connected to the sensor layer.
- Charger to the transmitter incl. cable.

Figure 1 The Heylo™ app, sensor layer and transmitter

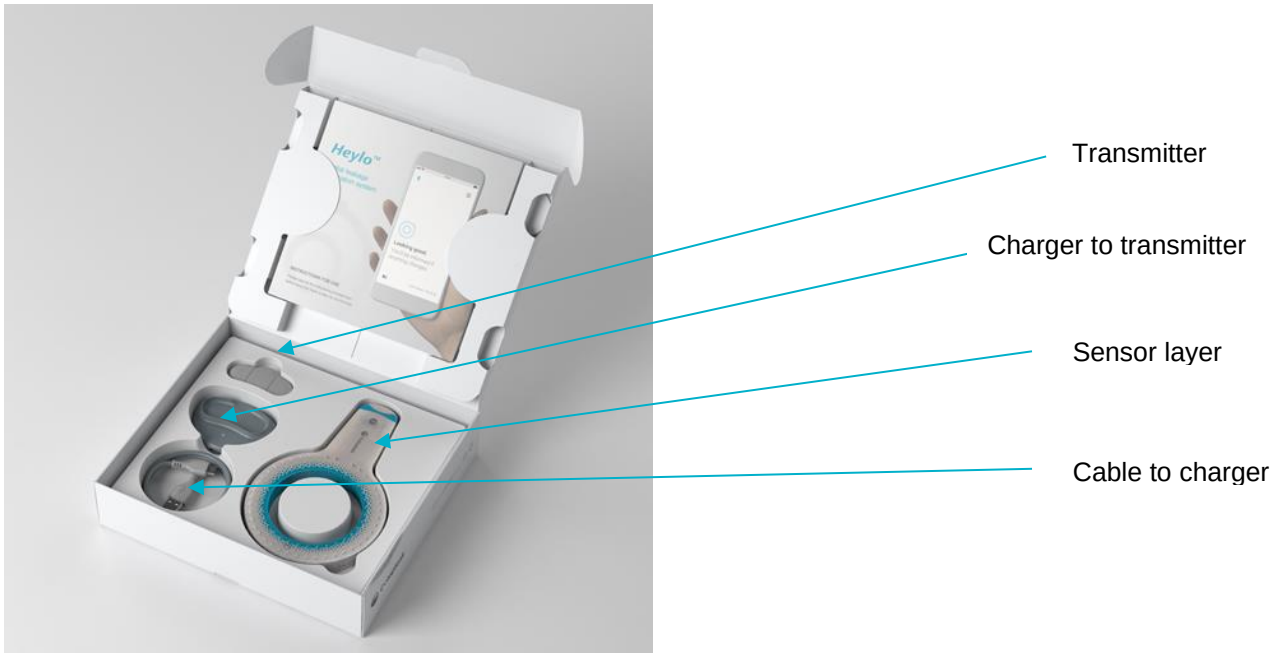


Figure 2 The investigational device - Heylo™

5.1.1 The Heylo™ app

A bespoke smartphone software application is installed on the user's smartphone. The app informs the user of any changes to the user's baseplate status, based on measurements obtained from the sensor layer and transmitter.

The app displays the status of the baseplate to the user via different app screens and pushes notifications to the smartphone home screen in case of relevant changes. The app can for instance inform the user of: "*No leakage detected (Looking good)*", "*There is a problem, it may be leakage*", "*The problem is spreading*", "*Bluetooth connection is lost*" etc. (Figure 3).

The Heylo™ app is delivered in local language. Thus, in this specific investigation the app was tested in German language. The app version tested in this clinical investigation was 2.3.0 for both Android and iOS.

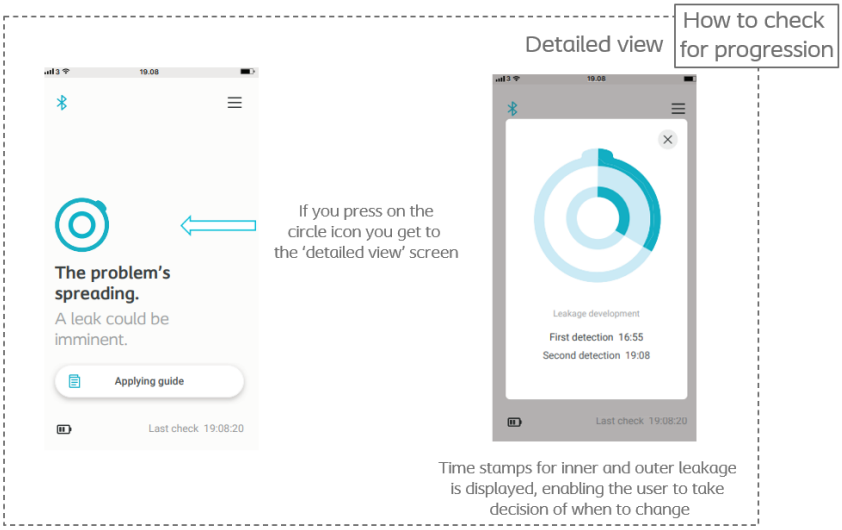


Figure 3 Example of status conveyed to a user

5.1.2 The sensor layer and the transmitter

The following sensor layer sizes were available in the study: ø40, ø50, ø60, ø70 and ø80, referring to the inside diameter (in mm) of the sensor layer (Figure 4A). The range of sensor layer sizes covers most ostomy product solutions on the market including distinct types of ostomy solutions (1-piece and 2-piece products and concave, convex and flat products) and brands (Coloplast, Hollister, ConvaTec etc.).
In this investigation the study nurse and the subject found the right size of sensor layer that best fitted the subject's stoma appliance together.

The colours on the sensor layer are only illustrative and illustrate the different sensor elements (Figure 4B):

- *Turquoise*: Outer leakage sensor. Detecting leakage close to the rim of the baseplate
- *Orange*: Wear sensor. Detecting moisture absorbed by the adhesive material
- *Purple*: 3 inner leakage sensors. Detecting leakage closest to the stoma

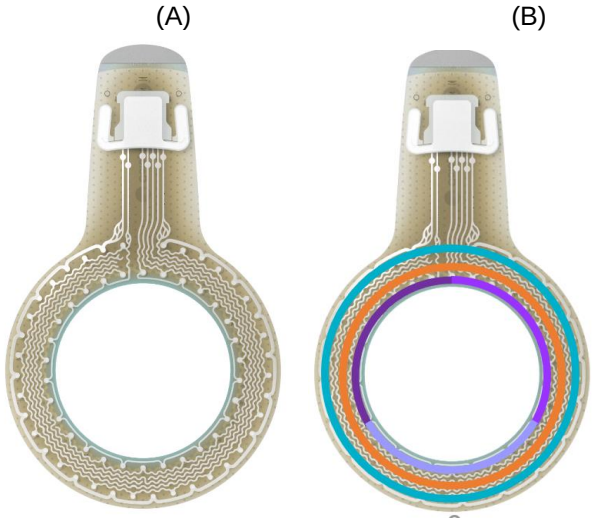


Figure 4 Sensor layer

Figure 5 Application of sensor layer onto baseplate

The transmitter is attached to a connector on the sensor layer and forms electrical low-voltage circuits to which voltage changes can be measured and used to calculate the state of the sensor layer. The status of the sensor layer units is sampled >1000 times an hour. An underlying algorithm and status flow decides which information to deliver to the user about the baseplate status.

The investigational device was used in combination with the ostomy solution usually used by the subject (Figure 5). Subjects were supplied with two transmitters (to secure continuity of use in the study set-up), a charger, a charger cable, and enough Heylo™ sensor layers to support their normal change routine.

5.1.3 Intended use of the investigational device in the clinical investigation

5.1.3.1 Intended purpose of the investigational device

Heylo™ is intended to be used together with an ostomy baseplate and bag, to detect and notify the user of the occurrence of output leakage under an ostomy baseplate.

5.1.3.2 Intended medical indication(s)

The product is indicated for users with an ostomy, mainly ileostomy and colostomy with liquid and/or mushy output. The product is to be used on intact skin.

5.1.3.3 Intended mode of action

The sensor layer must be applied underneath an ostomy baseplate that is then attached to intact peristomal skin. The sensor layer detects the occurrence of leakage underneath the ostomy baseplate and the user is notified of the leakage via a bespoke smartphone software application (The Heylo™ app).

5.1.3.4 Application

In the Instruction for Use (IFU), few warnings, cautions and pre-caution on how to use Heylo™ have been provided (See IFU (13, 14)). A physical IFU was delivered in the Heylo™ starter kit.

5.2 Description of standard of care (comparator product)

The comparator product in this study was standard of care, which was defined as the subject´s own ostomy product solutions (bags, baseplates, and accessories) including distinct types of ostomy product solutions (1-piece and 2-piece products with different baseplate shapes, including concave, convex and flat adhesive shape) and brands (Coloplast, Hollister, ConvaTec etc.). A typically changing pattern of an ostomy solution for people living with a stoma is from two times per day (1-piece users) up to once every 4th day (2-piece users), or even less frequent. In this investigation subjects were requested to follow their usual changing pattern.

5.3 Clinical Investigation Plan

The clinical investigation was carried out in accordance with Coloplast clinical investigation plan (CIP) VV-0336576 (15), study ID CP345 as amended and approved by all relevant parties.

No amendments were made.

5.3.1 Clinical investigation objective

The primary objective was to investigate whether Heylo™ can improve the *Emotional impact* domain score of the Ostomy Leak Impact tool (OLI), compared to standard of care after 8 weeks of product usage.

The secondary objective was to evaluate whether Heylo™ can improve participation in everyday- and social activities measured by the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) *Participation* domain (domain 6) compared to standard of care after 8 weeks of product usage.

5.3.2 Clinical investigation design

This clinical investigation was an open-label, randomized controlled cross-over trial comparing Heylo™ with standard of care. The investigation was planned to be conducted in Germany with up to 16 Coloplast Homecare nurses acting as study nurses. The principal investigator, who was responsible for conducting the investigation, delegated the study-specific procedures to the study nurses, who performed study visits at the subject's home or through remote virtual calls.

The expected duration of the total investigation period was 16 weeks (±6 days). Due to the cross-over study design each subject therefore tested Heylo™ for 8 weeks (±3 days) and standard of care for 8 weeks (±3 days), for allocation of treatment sequence see section 5.3.17 (Figure 6). Subjects used Heylo™ together with their own product (baseplate and bag) during the test period with Heylo™ and in the standard of care period subjects only used their own product (baseplate and bag), see Figure 6 for the design of the clinical investigation.

In both test periods subjects changed ostomy products as usual, according to their usual changing pattern. Subjects may have decided to change ostomy products earlier or later, based on a leakage notification.

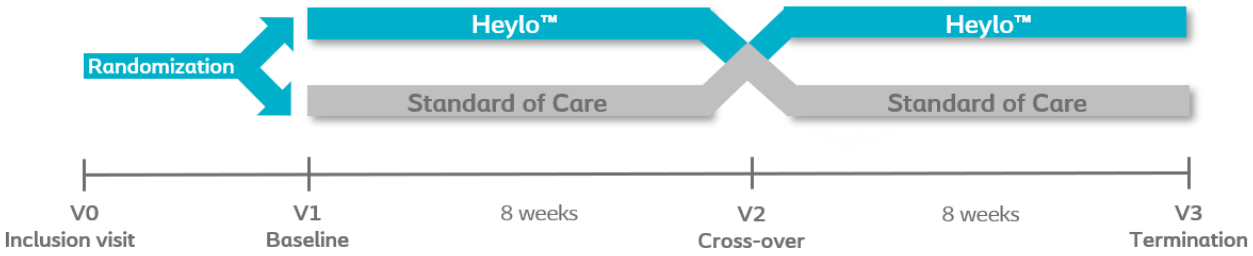


Figure 6 Design of the randomised controlled cross-over trial
All subjects who met the inclusion and exclusion criteria were randomised to one of two treatment sequences at the baseline visit (V1), with a cross-over after 8 weeks. Both sequences examined the investigational device (Heylo™) and the standard of care product. Sequence A: subjects start on Heylo™ and cross-over to standard of care and Sequence B: subjects start on standard of care cross-over to Heylo™.

In total, 144 subjects were planned to be enrolled and randomized (see section 5.3.20.8). Potential study participants being >18 years, with a colostomy or ileostomy were found through the Coloplast database and were contacted either by letter, e-mail, or phone as first contact. All subjects who were interested and found eligible as per study inclusion/exclusion criteria were consecutively enrolled into the investigation.

Each subject had an inclusion visit (V0), a baseline visit (V1), at which timepoint the baseline survey was conducted, and two test visits - V2 (8 weeks, period 1) and V3 (8 weeks, period 2). V3 was the termination visit unless a situation occurred where a subject terminated earlier than expected. If this was the case the subject had, as the last visit, the termination visit performed. For all patients, reason for discontinuation was obtained. All visits were carried out by the principal investigator, or delegate.

A follow-up call was scheduled 7 days ±2 days after visit 1 (V1) and visit 2 (V2) to ensure compliance with the provided product, the study procedures and assurance of subject's wellbeing. Additional calls/visits were performed if needed, assessed by the principal investigator or delegates, and registered as unscheduled visits.

Before scheduling V0 the subject was invited to an information meeting, where the principal investigator or the study nurse gave detailed information about the requirements and the content and what it involved to participate in the study. The information meeting was conducted as a phone call. It was possible to perform the inclusion visit (V0) and baseline visit (V1) on the same day. If the subject wanted to reconsider his/her participation at V0 after another oral review of the content of the study, the subject had the right to wait

minimum 24 hours before deciding on participation. If the subject hereafter decided to participate in the clinical study, another date for V0 and V1 was scheduled.

The study visits were conducted at the subject´s home or through remote virtual calls. For a detailed overview of the connection between visits and assessments see Table 1.

For a detailed overview of endpoints and assessments, see Appendix 12.8. At V1, V2 and V3 the principal investigator or delegate instructed the subjects in completing questionnaires including the OLI tool and the WHODAS 2.0. Every 2nd week the questionnaires and a notification were sent to the subject. A reminder to complete the questionnaires was sent again after 2 days. The OLI questionnaires and the questions about leakage episodes outside the baseplate were completed every 2nd week, whereas the WHODAS 2.0 questions were completed every 4th week. All questionnaires were completed using the electronic data-capturing system Smart -Trial (section 5.3.13).

Table 1 Relation between visits and assessments

	PERFORMED BY	INCLUSION VISIT	BASELINE VISIT	TEST VISIT	TEST VISIT AND TERMINATION VISIT	FOLLOW-UP CALL
VISIT	-	**V0**	**V1**	**V2**	**V3/TERMINATION VISIT**	**7 DAYS ±2 AFTER VISIT 1 AND VISIT 2**
WEEK	-	**WEEK 0**	**WEEK 0**	**WEEK 8**	**WEEK 16**	**WEEK 1 AND WEEK 9**
VISIT WINDOW	-	-	-	**±3 DAYS**	**±3 DAYS**	-
GENERAL						
Review of Subject information	Investigator	X				
Signed informed consent Form	Investigator	X				
Inclusion and allocation of subject number	Investigator	X				
Check of in- and exclusion criteria	Investigator	X				
Collect Baseline information	Investigator	X				
Insurance of subjects' wellbeing	Investigator	X	X	X	X	X
Insurance of subjects' compliance with the CIP	Investigator	X	X	X	X	X
Randomization	Investigator		X			
Instruction in the installation of the Heylo™ app and instruction in access to link with questionnaires	Investigator/Subject		X	X		
QUESTIONNAIRS						
Emotional impact (OLI)	Subject		X	X	X	
Coping and in control (OLI)	Subject		X	X	X	
Usual and social activities (OLI)	Subject		X	X	X	
WHODAS 2.0	Subject		X	X	X	
Feeling of security	Subject		X	X	X	
Leakage outside baseplate	Subject		X	X	X	

Change in current stoma product	Investigator			X	X	
Change in Heylo™ size	Investigator			X	X	
Remind subject to complete questionnaires between visits	Investigator		X	X		
PROCEDURES						
Instruction in the use of Heylo™ and SoC, according to IFU (if relevant according to randomization)	Investigator		X	X		
Measure size and provide Heylo™ (according to Randomization)	Investigator		X	X		
Collect unused Heylo™ devices (if relevant) and used charger and transmitter	Investigator			X	X	
Schedule follow-up call 7 days ±2 days after visit 1 and visit 2	Investigator		X	X		
Schedule next visit	Investigator		X	X		
Assess AEs/ADEs/SAEs/SADEs, Device Deficiencies and protocol deviations	Investigator		X	X	X	X
Complete eCRF	Investigator	X	X	X	X	X

5.3.3 Primary endpoint

The primary endpoint was the *Emotional impact* domain score (scale from 0 – 100, higher scores correspond to better emotional status) measured by the OLI tool evaluated after 8 weeks on Heylo™ or standard of care.

5.3.4 Secondary endpoint

The Secondary endpoint was the *Participation* domain score (Domain 6) (scale from 0 – 100; the lower scores corresponding with better participation in everyday- and social activities) measured by the WHODAS 2.0 (self-administrated version) evaluated after 8 weeks on Heylo™ or standard of care.

5.3.5 Exploratory endpoints

- *Usual and social activities* domain score (scale from 0-100) measured by the OLI tool evaluated after 8 weeks on Heylo™ or standard of care.
- *Coping and in control* domain score (scale from 0-100) measured by the OLI tool evaluated after 8 weeks on Heylo™ or standard of care.
- *Cognition* domain (Domain 1) score (scale from 0-100) measured by WHODAS 2.0 evaluated after 8 weeks on Heylo™ or standard of care.
- *Mobility* domain (Domain 2) score (scale from 0-100) measured by WHODAS 2.0 evaluated after 8 weeks on Heylo™ or standard of care.
- *Self-care* domain (Domain 3) score (scale from 0-100) measured by WHODAS 2.0 evaluated after 8 weeks on Heylo™ or standard of care.
- *Getting along* domain (Domain 4) score (scale from 0-100) measured by WHODAS 2.0 evaluated after 8 weeks on Heylo™ or standard of care.

- *Life activities* domain (Domain 5) score (scale from 0-100) measured by WHODAS 2.0 evaluated after 8 weeks on Heylo™ or standard of care.

- Feeling of security evaluated after 8 weeks on Heylo™ or standard of care. Question: "*How was the feeling of security while wearing the product?"* Answers: Very poor/Poor/Acceptable/Good/Very good.

5.3.5.1 Additional exploratory endpoints (described in the statistical analysis plan):

- Total WHODAS 2.0 score (mean of all 6 WHODAS 2.0 domain scores) (scale 0-100, the lower score the better) after 8 weeks on Heylo™ or standard of care.

- *Days affected by disability*, measured by three additional WHODAS 2.0 questions (Questions H1–H3) evaluated after 8 weeks on Heylo™ or standard of care.

5.3.6 Assessments and safety assessments:

- Leakage outside the baseplate, evaluated after 8 weeks on Heylo™ or standard of care. Question: "*Think back on the last 2 weeks; how many times have you experienced stoma effluent leakage outside the baseplate (e.g. onto clothes or bedsheets)?" (number).*

- Change in current ostomy solution, evaluated at the end of each test period. Question: "*Has there been any change in current ostomy solution during the test period?" (Yes/No) If yes, please add: Type (1pc/2pc), Kind (Flat, Convex, Concave), Brand (Coloplast, Hollister, Dansac, Salts, other)*

- Change in Heylo™ size, evaluated at the end of each test period. Question: "*Change of Heylo™ size needed?" (Yes/no), if yes, please provide the new size: 40 mm, 50 mm, 60 mm, 70 mm, 80 mm.*

- Adverse events (AEs)/device deficiencies (DDs)

5.3.7 Baseline information and potential compromising factors

The following baseline information and potential compromising factors were collected:

- Sex (male/female)
- Age (at time of enrolment (years)
- Height (cm)
- Weight (kg)
- Year of stoma creation (YYYY)
- Ostomy surgery within 3 months (yes/no)
- Reason for creation of the stoma (Crohn's disease/ulcerative colitis/ cancer/ Other)
- Stoma Type (ileostomy/colostomy)
- Temporary/permanent stoma
- Shape of the stoma (round/oval/irregular)
- Size of the stoma (widest diameter and height of stoma from skin)
- Information about the current stoma product: Type (1P/2P), Kind (Flat, convex, concave), Brand (Coloplast, ConvaTec, Hollister, Dansac, Salts, other)
- Working status (working, restricted duties, sick leave, unemployed/retired, student)

5.3.8 Methods for primary and secondary endpoints

For people living with a stoma, leakage and the related worry has both emotional, social and societal implications (3). Many people report that leakage outside the baseplate or the fear of leakage outside the baseplate made them stay at home, keep waking up at night, avoid contact with other people and/or affected them at work. The primary and secondary endpoints in this trial consisted of the *Emotional impact* domain and

Participation domain, and they are based on the validated measurement tools; OLI and WHODAS 2.0 respectively (16-18). These primary and secondary endpoints were chosen because they reflect relevant aspects of the emotional, social and societal burden faced by people experiencing leakage and worrying about leakage. Evaluation of these QoL domains as well as other questions related to coping and in control, impact on usual and social activities, feeling of security, getting along with people, and life activities may together establish possible positive care effects of using Heylo™ compared to standard of care.

5.3.8.1 The Ostomy Leak Impact tool

The OLI tool consists of 22-items interrogating the preceding 7 days (16). The questions in this tool are grouped into three domains with 4-10 questions in each domain (Appendix 12.9). The three domains concern the *Emotional impact* of having a stoma, *Usual and social activities*, as well as *Coping and in control*. Each domain has a score ranging from 0-100. A score of 100 equals no impact and a score of 0 represents full negative impact, thus a higher score means that subjects are less impacted by leakage in their daily life.

ICON Language Services conducted the translation from English to German and the linguistic validation of the OLS Tool (Appendix 12.9 (19)).

An important aspect for health care providers in evaluating the relevance of observed positive care effects of interventions in the clinic, is whether the change observed also translate into a clinically meaningful change for the patients and is reflected through a minimal clinically important difference (MCID) value. Different methods exist for the determination of an MCID-value (16, 20).

Through the development of the OLI tool, three individual MCID values for the different domain scores were determined using three different standard methods (16), and similar methods has been described for estimating MCID-values for the WHODAS 2.0 (21-26).

5.3.8.2 The World Health Organization Disability Assessment Schedule 2.0

The WHODAS 2.0 is a generic assessment tool developed by WHO to provide a standardised method for measuring health and disability across cultures (17, 18). The questionnaire consists of 36 items, assessing a wide range of abilities, each corresponding to an International Classification of Functioning, Disability and Health (ICF) code. Respondents are instructed to report any difficulties in activities or participation due to health conditions that have been encountered during the past 30 days.

In this study, the self-administered German version of the questionnaire was used (27). The questions are grouped into the following six domains (Domain 1-6) (For the English version of the WHODAS 2.0 see Appendix 12.10):

Domain 1: Cognition – assesses communication and thinking activities; specific areas assessed include concentration, remembering, problem-solving, learning, and communicating (items D1.1–D1.6)

Domain 2: Mobility – assesses activities such as standing, moving around inside the home, getting out of the home and walking a long distance (items D2.1–D2.5)

Domain 3: Self-care – assesses hygiene, dressing, eating & staying alone (items D3.1–D3.4)

Domain 4: Getting along – assesses interactions with other people and difficulties that might be encountered with this life domain due to a health condition; in this context, “other people” includes those known intimately or well (e.g. spouse or partner, family members or close friends) and those not known well (e.g. strangers) (items D4.1–D4.5)

Domain 5: Life activities – assesses difficulty with day-to-day activities (i.e. those that people do on most days, including those associated with domestic responsibilities, leisure, work & school (items D5.1–D5.8)

Domain 6: Participation – assesses social dimensions, such as community activities; barriers and hindrances in the world around the subject; and problems with other issues, such as maintaining personal dignity. The questions do not necessarily and solely refer to the ICF participation component as such, but also include various contextual (personal and environmental) factors affected by the health condition of the subject. (items D6.1–D6.8).

All 36 questions should be answered, except if the respondents are unemployed or no longer undergoing academic studies, then only 32 of the questions should be answered, leaving out items D5.5 through D5.8, since these items ask questions about functioning in work and/or studies.

Three additional questions (Questions H1–H3) summarise the extent to which the various difficulties respondents have encountered have affected their lives (Appendix 12.10) (17, 18).

5.3.9 Safety

See 12.7 for definitions of different AE types and DDs.

5.3.10 Ethical considerations

The clinical investigation was conducted in accordance with:

- Ethical principles that have their origin in the Declaration of Helsinki, 1964, Last amended at the 59th WMA General Assembly, Brazil, October 2013.
- MDD 93/42/EEC as amended by Directive 2007/47/EC (commonly known as the Medical Device Directive).
- MDR (EU) 2017/745
- ISO 14155:2020 “Clinical Investigation of medical devices for human subjects – Good clinical practices”.
- Any applicable regional or national regulations will be specified in the country-specific CIP.

The CIP and/or other relevant documents were submitted to the appropriate ethic committee (ECs) (application no. 201/2021). This clinical investigation was not initiated until the required approval from the EC was obtained. The Sponsor has notified the EC of the end of the clinical investigation.

5.3.11 Data quality assurance

To assure accurate, consistent and reliable data the Sponsor (Clinical Manager or a representative hereof) was responsible for:

1. Training of investigator and study personnel in the informed consent procedure, study procedures, how to use the products, complete the electronic Case Report Form (eCRF), how to report possible safety issues and in ISO 14155.
    - All training was documented by the site
2. Support during the recruitment process and conduct of the investigation
3. Remote monitoring of the data entered in the eCRF.

5.3.12 Monitoring Plan and Source data verification

All data collected were entered into the eCRF and the electronic data capturing (EDC) system performed edit checks to ensure that all fields were completed in the eCRF. The Monitor ensured that all queries were timely resolved.

Source data verification was performed to the extent it was possible. The Source Data Specification Form was completed at the initiation visit detailing the location of the source data for each data point. Where no source data (besides the eCRF) was available, the contents of the eCRF were monitored.

The Informed Consent Forms and AE/ Adverse Device Effect was 100% monitored for timely completeness.

Only the investigator, delegated site personnel and the sponsor representatives had access to all the CRFs. The subject had access to his/her CRF.

5.3.13 Data Management

Data was collected through an electronic data capturing (EDC) system; a secure, internet-based case report form. This system was used to record all subject information collected in the investigation for secure data tracking and centralised data monitoring ("remote monitoring") done by monitors, as defined in the CIP (15).

The EDC system used was Smart-Trial version 2021.4. The system is designed to be compliant with the FDA requirements of 21 CFR part 11. It is a validated data management system allowing only qualified and trained personnel to enter the system. The system has full audit trail and electronic signature.

The database consisted of two parts:

1) An electronic patient-reported outcome (ePRO), which collected subject questionnaire data directly from the subjects via a link sent out to the subject through email or SMS.
2) An electronic case report form (eCRF), which is used to collect data entered directly into the Smart-Trial by the site personnel.

The principal investigator, or delegate, entered data for each subject in the eCRF at the visit or immediately after. The eCRF made it possible to enter data right away when they were obtained. In case this was not possible the data was entered no later than 7 days after the visit/procedure.
If needed the investigator assisted the subject in completing the questionnaires.

Principal investigator, or delegate, at the clinical site performed primary data collection directly into the eCRF or drawn from source-document (i.e. medical records) reviews. The eCRF was completed continuously starting from the point of enrolling the subject to final follow-up.

The sponsor was responsible for the training of the investigator, or delegate, in the completion of the eCRF.

The eCRF was completed by the investigator, or delegate, who had signed the Site Personnel Signature and Delegation List and Clinical Investigation Training Log. It was the responsibility of the investigator to ensure that all measurements and observations were correctly noted in the eCRF.

All AEs were registered as described in the AE section (*Section 12.7*)

In the unforeseen situation, where the site could not establish a connection to the EDC system a paper CRF (pCRF) had been printed and supplied by the sponsor.

The investigator kept a separate list of the subjects' ID numbers, names and addresses in a locked room/cabinet. Only data referred to in the CIP (15) was recorded in the CRFs.

5.3.13.1 Database Management, Queries and Quality Control

The data management system had restricted role-based access control. The principal investigator or delegate was trained in the system before getting access. The training was completed before access to the investigation was granted.

The monitor, using his/her personal login information verified all critical data points against the source documents and issue electronic queries for the authorised clinical site personnel to respond, as defined in the CIP (15).

The principal investigator, using his/her personal login information signed each eCRF.

Automated, real-time access to the data enabled control of study compliance and safety assessments. Automated alerts (emails) were generated by the system to ensure full control and easier compliance with the CIP (15).

Critical quality control was performed by the sponsor's data management team and queries were issued where needed. Such queries were reviewed by the monitor and resolved by the site personnel.

At the end of the study, a formal data review meeting was performed before the database was locked (28). A full audit trail ensured, that each user's (site personnel, monitor, sponsor, data manager) access to and actions in the system has been tracked.

Data management and the final statistical analyses of all measurements described were carried out by Medical Affairs, Coloplast A/S.

5.3.13.2 Data Retention

The sponsor file will be archived for a minimum period of 10 years after the final clinical investigation report has been signed.

All investigation site documents will be archived for a minimum period of 10 years after the final clinical investigation report has been signed.

5.3.14 Subject population

5.3.14.1 Inclusion criteria

A subject was eligible for inclusion in the trial if all the following criteria applied (Table 2):

Table 2 Inclusion criteria

Inclusion criteria	To be included in the evaluation a subject must comply with the following inclusion criteria:	Justification for inclusion criteria
1.	**Has given written consent to participate by signing the Informed Consent Signature Form**	To meet the Helsinki declaration
2.	**Is at least 18 years of age and have full legal capacity**	To meet the Helsinki declaration
3.	**Has an ileostomy or colostomy with consistent liquid/mushy fecal output (5-7 Bristol scale*)**	The product is indicated for use with ileostomies and colostomies with liquid fecal output

Inclusion criteria	To be included in the evaluation a subject must comply with the following inclusion criteria:	Justification for inclusion criteria
	***See appendix 12.11**	
4.	**Is able to use one of the five investigational devices (i.e. Ø40, Ø50, Ø60, Ø70, Ø80 mm)**	The technical design of the device requires use of one of the five sizes
5.	**Has experienced leakage** under the baseplate at least three times within the last fourteen days. *****Leakage defined as output seeping under the baseplate" Appendix 12.12, Figure 2-5***	To ensure that the subjects have potential leakage which the sensor layer can react upon
6.	**Has worry about leakage to "some degree, high degree, or very high degree" (on a five-point Likert scale: Very low degree/Not at all, Low degree, some degree, High degree, very high degree)**	To ensure only subjects that worry about leakage are included
7.	**Is willing to refrain from use of ostomy paste**	The use of ostomy paste may influence system performance
8.	**Has a smartphone compatible with the Heylo™ application**	To answer the CRF questions and handle the Heylo™ applications, the subjects must have smartphones
9.	**Is able to follow study procedures for 4 months (assessed by investigator or delegate)**	To ensure a low drop-out rate

5.3.14.2 Exclusion criteria

A subject was not eligible for inclusion in the trial if any of the following criteria applied (Table 3):

Table 3 Exclusion criteria

Exclusion criteria	A subject is not allowed to participate in case he/she:	Justification for exclusion criteria:
1.	**Is participating in other clinical investigations or has previously participated in this investigation**	Other investigational guidelines/products may interfere with the investigational endpoints
2.	**Is pregnant or breastfeeding**	Even though the ingredients and the recipes have been approved for human beings, their effect on embryos, foetuses and infants are unknown
3.	**Has known hypersensitivity towards any of the products used in the investigation**	It is not ethical to include persons that know they are allergic to the products used in the investigation and it would also create bias, as these persons would give the product, they are allergic to a more negative rating and most likely also create an AE.
4.	**Is using/ has a pacemaker**	To protect the subjects from unnecessary harm, subjects with pacemakers are excluded

5.3.15 Pregnancy and breastfeeding

As specified in exclusion criteria number 2. pregnancy and breastfeeding were not allowed in this clinical investigation. All female subjects with childbearing potential (they have had at least one period during the last 12 months), had to confirm at V0 that they were not pregnant or breastfeeding. They were also informed that no pregnancy was allowed during the investigation.
If the subject became pregnant during the investigation, it was important, that the subject informed the Investigator/Investigator representative immediately. The principal investigator then considered whether she could continue in the investigation.

5.3.16 Subject Screening and randomisation failures

Subjects who signed the informed consent form, but failed to comply with inclusion or exclusion criteria, were considered screening failures.

5.3.17 Treatment allocation schedule

All subjects who met the inclusion and exclusion criteria were randomised to one of two treatment sequences at V1, with a cross-over after 8 weeks (V2). Both sequences examined the investigational device (Heylo™) and the standard of care product.

Treatment sequences:

Sequence A: Heylo™ cross-over to standard of care.
Sequence B: Standard of care cross-over to Heylo™.

A centralised randomisation was performed by Smart-Trial with random block sizes of 2, 4, 6, 8 or 10.
No stratification was applied by Smart-Trial, but it was post-hoc evaluated that 12 of the 31 subjects with stoma surgery less than 3 months were randomised to start with Heylo™ whereas 19 subjects were randomised to used Heylo™ in the second test period.

5.3.18 Concomitant medication/treatment

Concomitant medication/treatment was not registered in this study.

5.3.19 Duration of safety follow-up

AEs and DDs were assessed at all visits, planned and unplanned. The principal investigator ensured that adequate medical care was provided, during and after participation in the clinical investigation, if a subject experienced an AE. All ongoing Adverse Device Effects (ADEs), Serious AEs (SAEs), Serious Adverse Device Effects (SADEs) and DDs that could have led to an SAE at subject termination were followed according to the Risk-Benefit analysis (see section 7.7) and were followed until a resolution was addressed for a period of 2 months after subject termination. An ongoing AE at the subject termination visit was documented as the current status for the AE and was not followed up.

The subjects had to be informed of any new significant findings occurring during the clinical investigation, including the need for additional medical care that can be required, and the nature and possible cause of any AEs experienced.

The principal investigator or delegate had to provide the subject with the necessary instructions on the proper use, handling, storage and return of the investigational device when it was used or operated by the subject.

5.3.20 Statistical design, method, and analytical procedures

The primary objective was evaluated by analysing the primary endpoint, whereas the secondary objective was evaluated by analysing the secondary endpoint. The estimated treatment difference after a test period of 8 weeks should be statistically significant for at least one of the two endpoints, with the level of significance set at $p<0.025$ (see section 5.3.20.5 for adjustment for multiple testing). The analyses of the exploratory endpoints were used to further evaluate and explore the primary and secondary objectives and were all tested as 2-sided tests with the level of significance set at $p<0.05$.

All baseline measurements, endpoints and assessments were summarised by descriptive statistics and/or listed. Endpoints and assessments were summarised by product and time for evaluation, if relevant.

Descriptive statistics for continuous variables are presented with N, Mean, SD (standard deviation), Median, Min and Max, where n denotes the number of subjects contributing with non-missing data. For discrete variables, descriptive statistics are presented with N and percentage, where percentage is based on the total number of subjects/observations with non-missing data.

All statistical analyses are made with SAS version 9.4 (SAS Institute Inc., Cary, NC). For further information regarding the applied statistics, please see the Statistical Analysis Plan (SAP) – CP345 (29).

5.3.20.1 Definition of analysis populations

The intention-to-treat (ITT) population (full analysis set) constituted of all randomised subjects with valid informed consent who have been exposed to at least one product, with information on at least one endpoint. The safety population constituted of subjects who had given informed consent.

The ITT population and the safety population were defined at a formal data review meeting before database lock, see Data Evaluability Form for CP345 (28). The data manager, the clinical managers, the project manager and the statistician were involved in the classification of subjects. All Sponsor representatives, except the Data manager were blinded up until database lock.

All statistical analyses were based on the ITT population (full analysis set) whereas AEs and DDs were assessed based on the safety population. Invalid individual data points were omitted from the analysis even though the corresponding subject was part of the ITT population. Any exclusion of data points was documented before database lock (Data Evaluability Form for CP345 (28)).

A formal per-protocol (PP) population was not planned. Nevertheless, the protocol allowed for additional explorative analyses based on a subset of the ITT population, which has been performed when assessed to be relevant (see section 5.3.20.7).

5.3.20.2 Statistical analysis of the primary endpoint

The *Emotional impact* domain score (scale from 0-100) measured every 2nd week was analysed by a linear mixed model. The model included a fixed effect of product (standard of care, Heylo™), a fixed effect of time (2, 4, 6 and 8 weeks), a fixed interaction between product and time, a fixed period effect (test period 1 and 2) and a random effect of subject.

The difference between standard of care and Heylo™ at week 8 was estimated from the interaction between product and time as the contrast corresponding to week 8 (the primary comparison in the linear mixed model). The differences for weeks 2, 4 and 6 were similarly estimated and presented in figures and time trend evaluations.

The null hypothesis is taken to be “no difference in mean” between treatment groups after 8 weeks of treatment against the general alternative of “difference in mean” between the two groups.

H_0: $\mu_{Heylo,Week8} - \mu_{SoC,Week8} = 0$

H_1: $\mu_{Heylo,Week8} - \mu_{SoC,Week8} \neq 0$

The null hypothesis was tested based on a two-sided significance level of 2.5%

5.3.20.3 Statistical analysis of the secondary endpoint

Participation domain score (scale from 0-100) measured by WHODAS 2.0 was analysed by a similar model as the primary endpoint except that the questions were only filled out after 4 and 8 weeks in each test period.

The null hypothesis is taken to be “no difference in mean” between treatment groups after 8 weeks of treatment against the general alternative of “difference in mean” between the two groups.

H_0: $\mu_{Heylo,Week8} - \mu_{SoC,Week8} = 0$
H_1: $\mu_{Heylo,Week8} - \mu_{SoC,Week8} \neq 0$

The null hypothesis corresponding to the secondary endpoint was tested based on a two-sided significance level of 2.5% irrespective of whether the null hypothesis for the primary endpoint was rejected.

5.3.20.4 Statistical analysis of the exploratory endpoints

Impact on the *Usual and social activities* domain score (scale from 0-100) and impact on *Coping and in control* domain score (scale from 0-100) were all analysed by the same model as for the primary endpoint.

The remaining 5 domain scores based on the WHODAS 2.0 questionnaires (scale from 0-100) were all analysed by the same model as for the secondary endpoint.

For the remaining domain scores of the Ostomy Leak Impact tool and the WHODAS 2.0, the null hypothesis is likewise taken to be “no difference in mean” between treatment groups after 8 weeks of treatment against the general alternative of “difference in mean” between the two groups.

H_0: $\mu_{Heylo,Week8} - \mu_{SoC,Week8} = 0$
H_1: $\mu_{Heylo,Week8} - \mu_{SoC,Week8} \neq 0$

The feeling of security evaluated at the end of each test period (5-point Likert scale) was analysed by a generalised linear mixed model, namely a proportional odds model. The model included a fixed effect of product (standard of care, Heylo™), a fixed effect of period (test period 1 and 2) and a random subject effect.

The null hypothesis for feeling of security after 8 weeks of treatment is taken to be the odds ratio (OR) for the two treatment groups is equal to one against the general alternative of the odds ratio being different from one.

H_0: $OR_{Heylo/SoC(Week8)} = 1$
H_1: $OR_{Heylo/SoC(Week8)} \neq 1$

5.3.20.5 Adjustment for multiplicity

The results from the analyses of the primary and secondary endpoint were adjusted for multiple testing by a Bonferroni correction to keep a family-wise type 1 error of 5%. The Bonferroni correction for 2 endpoints corresponds to evaluating the difference between products as significant if the p-value is less than 0.025.

The applied testing strategy corresponds to a testing strategy with two primary endpoints. This strategy was chosen because a positive medical benefit on either the primary or secondary endpoint was considered sufficient.

For the analysis of the exploratory endpoints, no adjustment for multiple testing was applied.

5.3.20.6 Additional analyses and analysis of assessment

Analysis of the following assessment was pre-specified in SAP (29): Leakage outside baseplate (Question: "Think back on the last 2 weeks; how many times have you experienced stoma effluent leakage outside the baseplate (e.g. onto clothes or bedsheets)?"). The question has been asked every 2nd week.

Due to the potential of many reporting a low number of episodes of leakage within the 2-week period a Poisson distribution was used for modelling data instead of approximating to a normal distribution. The comparison between the mean numbers of leakages after 8 weeks for the two products was performed by a generalised linear mixed model. The model included a fixed effect of product (standard of care or Heylo™), a fixed effect of time (2, 4, 6 and 8 weeks), a fixed interaction between product and time, a fixed period effect and a random effect of subject. By using a negative binomial distribution, we allowed for over-dispersion of the Poisson parameter, if present. The relative risk for the 2 treatment groups at week 8 was estimated from the interaction term between product and time. From this interaction the contrast between treatment groups after 8 weeks product usage can be derived.

The null hypothesis for leakage outside baseplate after 8 weeks of treatment is taken to be that the relative risk (RR) for the two treatment groups is equal to one against the general alternative of the relative risk being different from one.

H_0: $RR_{Heylo/SoC(Week8)} = 1$
H_1: $RR_{Heylo/SoC(Week8)} \neq 1$

Further we pre-specified in the SAP (29) that the three additional WHODAS 2.0 questions (H1, H2 and H3) were analysed by the same model as described for the secondary endpoint. The answers to the questions were an exact number of days with disability in the range between 0 and 30. The total summary score of WHODAS 2.0 (scale from 0-100) was also analysed by the same model as described for the secondary endpoint.

5.3.20.7 Post-hoc sensitivity analyses

5.3.20.7.1 Analyses of the primary and secondary endpoints based on subjects with observations in both periods

As a post-hoc sensitivity analysis, the analysis of the primary endpoint and secondary endpoint was repeated including only subjects that had observations in both treatment periods at week 8. Hence, the impact of imputation of possible missing values, as done by the linear mixed model, could be evaluated.

5.3.20.7.2 Sub-group analyses

To evaluate the effect of type of ostomy solution (2-piece or 1-piece), the effect of the shape of the ostomy solution (convex vs flat/concave) and the effect of having a newly formed stoma (new patient discharged (NPD) defined as having ostomy surgery within the past 3 months) on the primary endpoint, the secondary endpoint and on the number of leakages outside the baseplate, the following effects were added to the analyses: Type (1-piece, 2-piece), shape (convex, not convex), NPD (Yes, No) and the interactions between type and product (Heylo™ vs. standard of care), shape and product as well as the interaction between product and NPD. It was then tested if the effects in the model are significantly different from zero. If the effects of the interaction with product are significant it indicates that the effect of Heylo™ is different for e.g., type of ostomy solution, shape of the ostomy solution or NPD/not NPD. If the effect of the interactions are not significantly different from zero, it might be because, there is no difference, the effect is very small, or the sample size is too small to find a relevant difference.

Sub-group analyses for the following endpoints were performed: the primary endpoint, the secondary endpoint and the number of leakages outside the baseplate. The statistical analyses were the same as described above.

Sub-groups consisted of only NPD subjects, only subjects with a temporary stoma and finally NPD subjects including those with a temporary stoma.

Sub-group analyses for sex (male/female), age groups (age 65 years or below/above 65 years), type of product (1-piece/2-piece), shape of product (convex/flat + concave), type of stoma (ileostomy/colostomy) were performed for the primary and secondary endpoints.

5.3.20.7.3 Responder analysis with absolute thresholds

As an additional sensitivity analysis, responder analyses for the primary and secondary endpoints were performed. Both the primary and secondary endpoints are QoL endpoints that are measured on continuous scales. These validated tools do not hold diagnostic thresholds, as for example is the case for the Beck Depression Inventory (30). Thus, a threshold for responder success would need to be decided based upon relevant literature.

For the primary endpoint, the *Emotional Impact* domain score (Scale ranging from 0 to 100), we have inspected the validation report for the Ostomy Leak Impact tool (31), and a global survey (Ostomy Life Study 2019 (3), conducted by Coloplast) that used the same tool to assess the impact of leakage across 17 countries. From the validation report of the tool, the population means of the *Emotional Impact* domain score for the United Kingdom and France were 67.0 and 66.9, respectively, and the mean for the 'global' population from the Ostomy Life Study 2019 was 67.7 (derived from the dataset presented in (3)). The populations in the Ostomy Leak Impact validation and in the Ostomy Life Study 2019 are representative of the general stoma population (32-34% had experienced leakage underneath the baseplate in the previous week in both studies (31, 32)), whilst the population included in the CP345 trial was struggling with leakage and worry about leakage (inclusion criteria).
Thus, the threshold for responder success was decided upon as values ≥67 for the *Emotional Impact* domain score.

For the secondary endpoint, the *Participation domain* of WHODAS 2.0, we have assessed that subjects who have *No problems* or are only *Mildly disabled* are considered a responder success and those who are *Moderately* to *Severely/Extremely* disabled are considered responder failures. The definitions of thresholds for the degree of disability are based on the publication from Lee et al. (33), with the following definitions: Score from 0-4 (*No problems*), score from 5-24 (*Mild disability*), score from 25-49 (*Moderate disability*) and a score from 50-100 (*Severe/Extreme disability*).

The responses (success/failure) based on the primary and secondary endpoints were analysed by a generalized linear mixed model that included a fixed effect of product (standard of care, Heylo™), a fixed effect of time, a fixed interaction between product and time, a fixed period effect (test period 1 and 2) and a random effect of subject. The response was assumed to be binomially distributed, and a logit was used as link function.

The marginal proportion was estimated for both endpoints. Furthermore, the odds ratio for the two groups (Heylo™ and standard of care) with 95% confidence interval was estimated at week 8 from the interaction between product and time as the contrast corresponding to week 8. The hypothesis that the odds ratio was equal to one (equal treatment effect after 8 weeks of treatment) was tested against the alternative hypothesis that the odds ratio was different from one (different treatment effect).

5.3.20.7.4 Responder summary based on changes from baseline

The proportions of patients who have deteriorated, improved or stayed unchanged compared to their exact baseline level have been summarized for all follow-up timepoints. Moreover, Sankey diagrams have been prepared to visualize the movement of responders over time.

5.3.20.7.5 Comparison of baseline characteristics of drop-outs and study completers

Summary statistics comparing baseline characteristics of the drop-out population by study groups (Heylo™ vs. standard of care), as well as characteristics of the completers were performed.

5.3.20.7.6 Analyses of the primary and secondary endpoints for all randomized subjects based on multiple imputation

The repeated mixed measure model (MMRM) analyses performed above did not include data from all randomized subjects, as post-baseline observations are a requirement for these analyses. To allow for all randomized subjects to contribute to the analyses, two different multiple imputation methods have been applied.

In one of them it was assumed that data was missing at random. In this way it resembled the prespecified repeated mixed measures model but allowed inclusion of randomized subjects where only the baseline values were available. First, within each sequence and separately for each period, intermittent data was imputed using a Markov Chain Monte Carlo method (MCMC), to obtain a monotone missing data pattern. 100 imputations were performed. For each of the 100 datasets, within each sequence and period, missing observations were imputed sequentially, including previous observations/imputed values as well as baseline as covariates. Each of the 100 datasets now had either observed or imputed week 8 assessments for all randomized subjects in both periods. For each dataset, an analysis was performed using a Mixed Model with the week 8 assessment as the response variable, treatment (Heylo™/SoC) and period (1/2) as fixed effects, and a random effect of subject. In another analysis, baseline, and the interaction between baseline and period were also included as covariates. The estimates and corresponding standard errors from the 100 datasets were pooled to one estimate and associated standard error using Rubin's rule. From these pooled estimates, the 95% confidence interval for the week 8 treatment contrast, and the associated p-value, were calculated.

The other method was a reference-based imputation method where it was assumed that missing observations (except for intermediate missing observations) in the period where subjects were randomized to Heylo™ were missing *not* at random. The reference based multiple imputation was performed similar to the one above assuming that data was missing at random. The difference was, how data was imputed after the monotone missing data had been obtained. Instead of imputing data within each sequence assuming that data was missing at random, the models used to impute the missing data for subjects randomised to Heylo™, as well as for subjects randomised to standard of care, were fitted only based on data from subjects randomized to standard of care in the corresponding period.

For a detailed description of both imputation methods see Appendix 12.17.

5.3.20.8 Sample size

The *Emotional impact* domain score (scale from 0-100) was measured every 2nd week and all measurements were part of the primary analysis, whereas the primary comparison was evaluated at the end of each test period (after 8 weeks).

The secondary endpoint WHODAS 2.0 *participation* (domain 6) was considered equally relevant for the overall aim of the investigation. However, since the use of this tool in ostomy care has only been sparsely investigated, sample size calculation was performed based on the primary endpoint only using a Bonferroni corrected significance level of 2.5%.

The sample size calculation for the primary endpoint was based on a simplified model (paired 2-sided *t* test). It was assumed that the total standard deviation of the primary endpoint was 20.6 and that the total variation 20.6^2 was divided so that the residual variation was 14.4^2 (based on data from the previous CP308 investigation (9)).

On the basis of the above assumptions, and considering that the true difference between the 2 treatments (with and without Heylo™) was at least 6 on the *Emotional impact* domain score (minimum clinically important difference is in the range of 5.4-10.4 according to the validation of the OLI tool) (16) a total of 108 participants should answer the questionnaires at the end of each test period to ensure a power of 81%. Taking a potential dropout of 25% into account, it was recommended to enroll 144 participants in the study.

5.3.20.9 Level of significance and power

A two-sided significance level of 5% was applied. For adjustment for multiple testing see section 5.3.20.5 and for a description of the power see the sample size section 5.3.20.8.

5.3.20.10 Pass/fail criteria

The purpose of the investigation was fulfilled if a statistically significant improved mean difference in the primary or secondary endpoint with a significance level of $p<0.025$ was obtained after 8 weeks usage of the investigational device compared to standard of care. The applied testing strategy corresponds to a testing strategy with two primary endpoints. This strategy was chosen because a positive medical benefit seen on either of the primary or secondary endpoint was considered sufficient.

5.3.20.11 Interim analysis

There was no planned interim analysis in this investigation.

5.3.20.12 Statistical reason for termination of investigation

There was no interim analysis and therefore no reason to terminate the investigation based on statistical considerations.

5.3.20.13 Deviation(s) from statistical design, method, or analytical procedures

Any deviations from the statistical plan have been documented in this clinical investigation report.

6 Results

6.1 Investigation period

Initiation date (First subject in)	13th of January 2022
Completed date (Last subject out)	4th of November 2022

6.2 Disposition of subjects and investigational device

This clinical investigation was conducted in Germany with 13 Coloplast Homecare nurses (out of up to 16 originally planned) acting as study nurses. Therefore, in this clinical investigation there were no regular Sites. The principal investigator, who was responsible for conducting the investigation, delegated the study specific procedures to the study nurses, who performed all study visits at the subject's home or by remote calls.

In all, a total of n=840 potential study participants were identified through the Coloplast database and contacted. Out of these n=563 (67%) showed interest in study participation and they were further screened for eligibility. A total of n=144 (26%) subjects were found eligible and were consecutively recruited and randomised (Figure 7). Table 4 presents the number of subjects recruited to the investigation, distributed between the 13 study nurses.

Table 4 Number of subjects recruited per study nurse

Country	Study nurse	No of subjects	% of subjects	No of non-completers
DE	Study nurse 1	5	3	1
DE	Study nurse 2	40	28	8
DE	Study nurse 3	4	3	0
DE	Study nurse 4	3	2	0
DE	Study nurse 5	11	8	1
DE	Study nurse 6	8	6	0
DE	Study nurse 7	9	6	0
DE	Study nurse 8	5	3	3
DE	Study nurse 9	7	5	0
DE	Study nurse 10	15	10	1
DE	Study nurse 11	10	7	2
DE	Study nurse 12	10	7	3
DE	Study nurse 13	17	12	1
Total		144	100	20

Table 5 Overview of safety and intention-to-treat population

	Number of subjects
Safety population	144
Intention-to-treat (ITT)	139

All enrolled (n=144) subjects were included as the safety population (Table 5). In all, 20 subjects did not complete the investigation as planned. A total of 5 subjects were omitted from the ITT population (full analysis set), as only baseline data, which were not part of the endpoints, were collected for these subjects. Furthermore, 15 subjects from the ITT population did not complete the investigation as planned; see Figure 7 for reasons. All data obtained from these subjects until the time of drop-out was included in the pre-specified statistical analyses. A more detailed description of drop-outs is provided in section 6.9.2 and section 6.9.3.

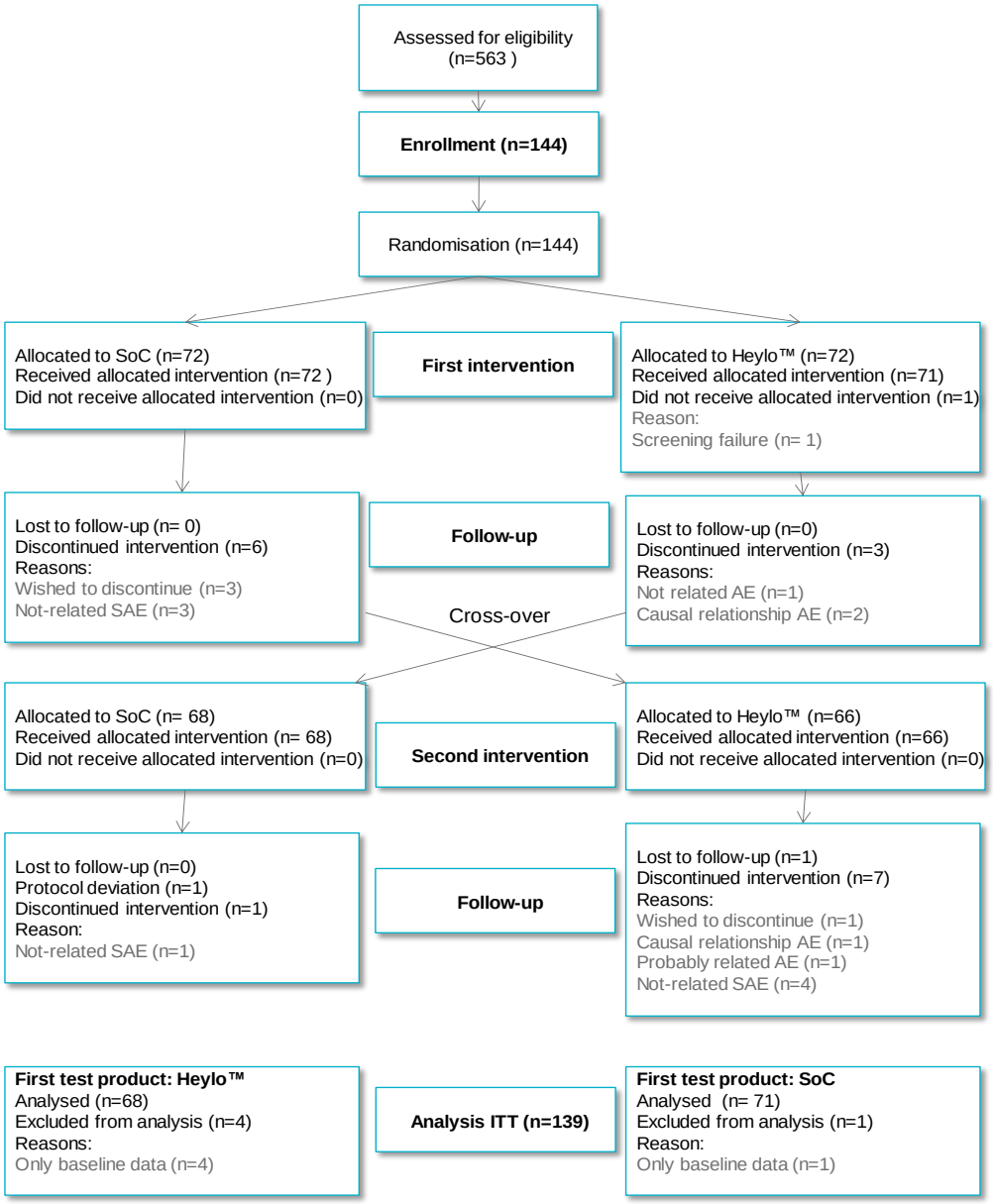


Figure 7 Subject disposition tree.

The intention-to-treat (ITT) population (full analysis set) was constituted of all randomised subjects with valid informed consent who have been exposed to at least one product, with information on at least one endpoint. The safety population constitutes of subjects who have given informed consent. A total of 5 subjects were omitted from the ITT population (full analysis set), as only baseline data, which were not part of the endpoints, were collected for these subjects. Abbreviations: SoC, Standard of care; SAE, Serious adverse event; AE, Adverse event.

6.3 Demographics

Table 6 displays the baseline and stoma characteristics of the ITT (full analysis set) population (n=139) and for the two test sequences (A and B). In all, 88 participants (63%) had an ileostomy and 51 (37%) had a colostomy. The population consisted of 71 females (51%) and 68 males (49%). The mean age was 50.7 years (SD= 13.5; range 18 to 81 years), with 12% of the population >65 years. Approximately half the participants were unemployed or retired, a quarter of the participants were working, while the remaining were on restricted duties (2%), sick leave (16%) or students (2%) (Table 6).

Subjects on average had their stoma for 6.1 years (SD=8.2; range 0 to 35 years;). A total of 31 subjects (22%) were new stoma carriers who had had their stoma for less than 3 months and 108 subjects (78%) were experienced stoma carriers who had had their stomas for more than 3 months (Table 6). In all, 21 out of the 36 subjects with a temporary stoma were new stoma users.

Table 6 Baseline characteristics

Baseline and stoma characteristics of the ITT population (full analysis set) and for subjects randomized to Test sequence A (start on Heylo™ with cross-over to standard of care (SoC)) and Test sequence B (start on standard of care with cross-over to Heylo™), respectively.

	Total ITT population n=139	**Test sequence A Heylo → SoC n=68**	**Test sequence B SoC → Heylo n=71**
Age in years; Mean (SD)	50.7 (13.5)	48.8 (13.5)	52.6 (13.2)
Min; Max	18; 81	18; 75	22; 81
Median	53.0	49.0	55.0
Adults 18-65 years; n (%)	123 (88%)	62 (92%)	61 (85.9)
Adults >65 years; n (%)	16 (12%)	6 (8.8)	10 (14.1)
Sex; females / males; n (%)	71 (51%) / 68 (49%)	38 (56%) / 30 (44%)	33 (46%) / 38 (54%)
Weight in kg; Mean (SD)	77.9 (18.6)	79.5 (18.8)	76.3 (18.4)
Height in cm; Mean (SD)	173.0 (8.2)	173.6 (8.2)	172.4 (8.3)
BMI; Mean (SD)	25.9 (5.5)	26.3 (5.6)	25.6 (5.4)
Working status			
Working; n (%)	37 (27%)	18 (26%)	19 (27%)
Restricted duties; n (%)	3 (2%)	2 (3%)	1 (1%)
Sick leave; n (%)	22 (16%)	11 (16%)	11 (15%)
Unemployed/retired; n (%)	74 (53%)	36 (53%)	38 (54%)
Student; n (%)	3 (2%)	1 (1%)	2 (3%)
Stoma characteristics at baseline			
Stoma age (years); Mean (SD, Min; Max)	6.1 (8.2, 0; 35)	6.0 (8.0, 0; 30)	6.3 (8.5, 0; 35)
Stoma surgery within the last three months; yes/no; n (%)	31 (22%) / 108 (78%)	12 (18%) / 56 (82%)	19 (27%) / 52 (73%)
Type of stoma; ileostomy / colostomy; n (%)	88 (63%) / 51 (37%)	44 (65%) / 24 (35%)	44 (62%) / 27 (38%)
Stoma sub-type; permanent / temporary; n (%)	103 (74%) / 36 (26%)	51 (75%) / 17 (25%)	52 (73%) / 19 (27%)
Shape of stoma; irregular / oval / round; n (%)	2 (1%) / 33 (24%) / 104 (75%)	1 (1%) / 14 (21%) / 53 (78%)	1 (1%) / 19 (27%) / 51 (72%)
Stoma diameter in mm; Mean (SD)	29.3 (7.7)	29.1 (7.7)	29.5 (7.7)
Stoma height in mm; Mean (SD)	21.9 (13.1)	21.1 (13.2)	22.6 (13.1)
Baseline measurements			
Ostomy Leak Impact			
Emotional Impact; Mean (SD)	51.7 (27.0)	49.4 (27.4)	53.9 (26.5)
Usual and Social Activities; Mean (SD)	62.2 (24.7)	62.1 (24.5)	62.3 (25.3)
Coping and in Control; Mean (SD)	56.3 (30.8)	54.0 (31.6)	58.5 (30.0)
WHODAS 2.0			
Domain 1 (Cognition); Mean (SD)	20.0 (22.5)	22.0 (23.1)	18.1 (21.8)
Domain 2 (Mobility); Mean (SD)	25.4 (26.1)	24.8 (26.7)	25.9 (25.8)
Domain 3 (Self-care); Mean (SD)	16.1 (24.3)	12.9 (21.8)	19.2 (26.2)
Domain 4 (Getting along); Mean (SD)	24.5 (25.1)	25.2 (27.4)	23.7 (22.9)
Domain 5 (Life activities); Mean (SD)	35.3 (30.8)	32.7 (29.1)	37.8 (32.3)
Domain 6 (Participation); Mean (SD)	39.2 (24.4)	40.7 (24.7)	37.8 (24.2)
Total score; Mean (SD)	26.7 (20.9)	26.4 (21.1)	27.1 (20.9)
Leakage onto clothes (last two weeks); Mean (SD)	3.0 (3.7)	3.3 (3.1)	2.8 (4.2)
Feeling of Security			
Very poor; n (%)	3 (2%)	2 (3%)	1 (1%)
Poor; n (%)	19 (14%)	12 (18%)	7 (10%)
Acceptable; n (%)	53 (38%)	24 (35%)	29 (41%)
Good; n (%)	47 (34%)	23 (34%)	24 (34%)
Very good; n (%)	17 (12%)	7 (10%)	10 (14%)

Most subjects used ostomy solutions manufactured by Coloplast (96.4%) as their regular solution, the remaining used ostomy solutions by other companies (ConvaTec, Hollister or other) (Table 7).

Table 7 Brand of ostomy solution

Ostomy solution brand	ITT population n=139 n (%)
Coloplast	134 (96.4%)
ConvaTec	2 (1.4%)
Hollister	2 (1.4%)
Dansac	0 (0%)
Salts	0 (0%)
Other	1 (0.7%)

Most subjects used a convex baseplate (n=88, 63%) while the remaining used a flat baseplate (n=34, 25%) or a concave baseplate (n=17, 12%). Also, there was an equal distribution of subjects who used 2-piece ostomy solutions (n=78, 56%) and subjects who used 1-piece solutions (n=61, 44%).

During the investigation, 4 subjects changed ostomy solution; two subjects changed solution in the first test period (one changed from a convex baseplate to a flat and one changed from a flat baseplate to a convex) and two subjects changed ostomy solution during the second test period (both changed from a convex baseplate to a flat).

Heylo™ was delivered in 5 different sizes. Table 8 displays the distribution of Heylo™ sizes at baseline in the ITT population. During the investigation, 4 subjects increased Heylo™ size, and 3 subjects decreased in Heylo™ size.

Table 8 Distribution of Heylo™ sizes at baseline in the ITT population

Heylo™ size	n (%)
40 mm	12 (9%)
50 mm	29 (21%)
60 mm	68 (49%)
70 mm	24 (17%)
80 mm	2 (1%)

n=4 (3%) had missing information regarding Heylo™ size at baseline.

Table 9 shows the underlying reasons for stoma formation of the study population as categorised by WHO-ART System Organ Class (SOC) codes. Of the ITT population, the main reason for having an ostomy was due to Crohn´s disease (34%), while 31% of the subjects had their ostomy because of colorectal cancer (Table 9). Ulcerative Colitis and "Other" were the reason for having an ostomy for 14% and 22% of the subjects, respectively.

Table 9 Reason for having an ostomy (ITT population)

Cause of ostomy	Total (n=139)
Crohn´s disease, n (%)	47 (34%)
Colorectal cancer, n (%)	43 (31%)
Ulcerative Colitis, n (%)	19 (14%)
Other, n (%)	30 (22%)

6.4 Clinical Investigation Plan compliance

The investigator was not allowed to deviate from the CIP unless under emergency circumstances and to protect the rights, safety, and well-being of the subject(s). No deviations affecting the scientific aspect of the investigation, or the safety of the subject(s) were observed.

During the database lock meeting, handling of deviations, missing data and allocation to ITT were dealt with. This is documented in the Data Evaluability Form for CP345 (28).

A full list of the deviations listed per subject can be found in the Statistical Analysis Report – Listings (34) and in the two note-to-files (35, 36).

6.5 Analysis

In general, all endpoints are analysed by models that include the investigational devices (product), time of evaluation, and the interaction between the two as well as the test period as fixed effects (dependent variables) and subject as a random effect (section 5.3.20). The primary comparison between the devices is performed after 8 weeks of use. Post-hoc analyses have been performed with all randomized subjects using two imputation methods (section 5.3.20.7.6).

6.6 Primary endpoint

The primary endpoint of this investigation was the *Emotional impact* domain score (scale from 0 – 100, higher scores correspond to better emotional status) measured by the OLI tool (section 5.3.8.1) evaluated after a test period of 8 weeks on Heylo™ or standard of care.

A significant improvement in mean *Emotional impact* score was found with Heylo™ compared to standard of care ($p < 0.001$) (Figure 8). The score increased from 62.0 to 73.4 (LS mean difference 11.4, 95%CI: 7.8; 15.0) with Heylo™ compared to standard of care.

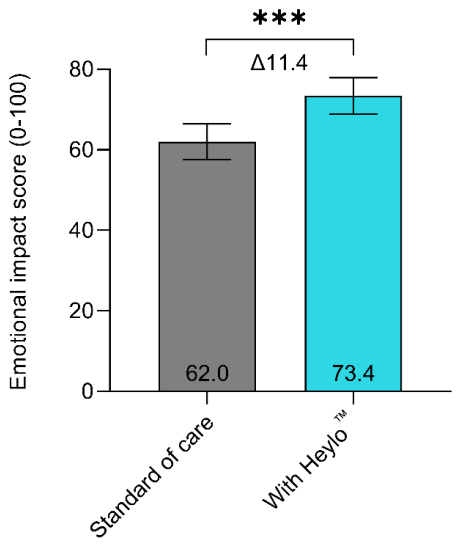


Figure 8 *Emotional impact* domain score.

LS mean values for the *Emotional impact* domain score and a P-value for comparison of Heylo™ and standard of care are displayed (n=139), obtained by linear mixed model as described in section 5.3.20.2. Error bars represent the 95% confidence intervals of the LS means. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

For summary statistics, a box plot of the raw data and plots for model diagnostics of the residuals from the general linear mixed model see Appendix 12.13 – 12.15.

6.7 Secondary endpoint

The secondary endpoint of this investigation was the *Participation* domain score (scale from 0 – 100, the lower scores corresponding with better participation in everyday- and social activities) measured by WHODAS 2.0 (section 5.3.8.2), evaluated after a test period of 8 weeks on Heylo™ or standard of care.

A significant improvement in mean *Participation* domain score was found with Heylo™ compared to standard of care (p = 0.001) (Figure 9). The score decreased from 37.1 to 33.00 with Heylo™ compared to standard of care (LS mean difference -4.2, 95% CI: -6.7; -1.6).

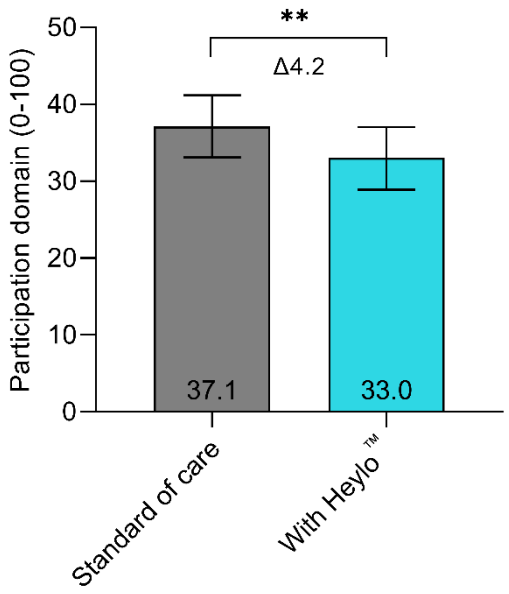


Figure 9 *Participation* domain score.

LS mean values for *Participation* domain score and a P-value for comparison of Heylo™ and standard of care are displayed (n=139) obtained by linear mixed model as described in section 5.3.20.3. Error bars represent the 95% confidence intervals of the LS means. ∗*P* < 0.05, ∗∗*P* < 0.01, ∗∗∗*P* < 0.001.

For summary statistics and a box plot for the raw data together with plots to check for normal-distribution of the residuals from the general linear mixed model see Appendix 12.13 – 12.15.

6.8 Exploratory endpoints

6.8.1 OLI scores for the two domains Usual and social activities and Coping and in control evaluated at the end of each test period.

The mean score for the two OLI domains *Usual and social activities* and *Coping and in control* improved significantly after 8 weeks with Heylo™ compared to standard of care ($p < 0.001$, respectively) (scale from 0 – 100, the higher score the better) (Figure 10). The mean score for the *Usual and social activity* domain increased from 73.4 to 81.1 (LS mean difference 7.7 95% CI: 3.5; 11.9) and the mean score for the *Coping and in control* domain increased from 65.9 to 73.7 (LS mean difference 7.8, 95% CI: 3.8; 11.7).

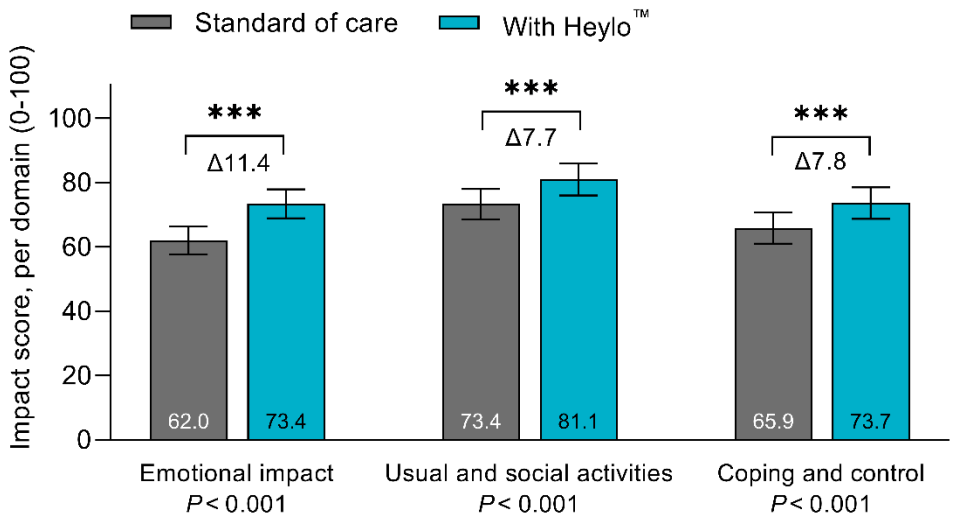


Figure 10 Ostomy leak impact score (OLI) per domain.

LS mean values for OLI scores per domain for Heylo™ and standard of care are displayed (n=139). Error bars represent the 95% confidence intervals of the LS means. ∗$P < 0.05$, ∗∗$P < 0.01$, ∗∗∗$P < 0.001$.

6.8.2 WHODAS 2.0 domain (Domain 1-6) scores and total WHODAS 2.0 score, evaluated at the end of each test period:

All six WHODAS 2.0 domain (Domain 1-6) mean scores and the mean total WHODAS 2.0 score improved significantly, except for the *Self-care* domain score (scale from 0 – 100, the lower score the better) (Figure 11):

- Domain 1: The mean *Cognition* domain score improved significantly after 8 weeks on Heylo™ compared to standard of care from 25.0 to 21.7 (LS mean difference -3.4, 95% CI: -6.1; -0.6, p=0.018).
- Domain 2: The mean *Mobility* domain score improved significantly after 8 weeks on Heylo™ compared to standard of care from 29.6 to 26.0 (LS mean difference -3.6, 95% CI: -6.5; -0.7, p=0.014).
- Domain 3: The mean *Self-care* domain score improved, but not significantly after 8 weeks on Heylo™ compared to standard of care from 19.2 to 17.2 (LS mean difference -2.1, 95% CI: -4.6; 0.5, p=0.120).
- Domain 4: The *Getting along* domain score (Domain 4) improved significantly after 8 weeks on Heylo™ compared to standard of care from 30.4 to 27.2 (LS mean difference -3.2, 95% CI -5.9; -0.5, p=0.020).
- Domain 5: The mean *Life activities* domain score improved significantly after 8 weeks on Heylo™ compared to standard of care from 38.9 to 34.1 (LS mean difference -4.8, 95% CI -8.3; -1.25, p=0.008).
- Domain 6: For specific results on the *Participation* domain score see section 6.7 Secondary endpoint.

- The mean total WHODAS 2.0 score (mean of all 6 domain scores) improved significantly after 8 weeks on Heylo™ compared to standard of care from 30.0 to 26.5 (LS mean difference -3.5, 95% CI: -5.5; -1.5, p<0.001).

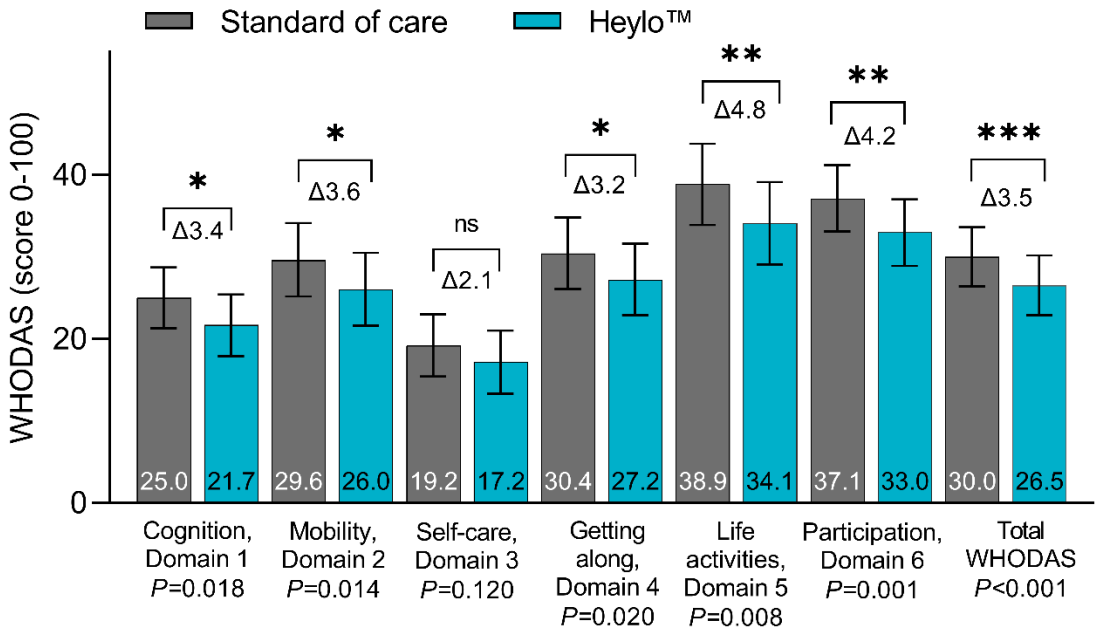


Figure 11 WHODAS 2.0 domain (D1-D6) scores and total WHODAS 2.0 score
Data is presented as LS means and error bars represent the 95% confidence intervals of the LS means (n=139).
*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

6.8.3 Days affected by disability, measured by three additional WHODAS questions (Questions H1–H3).

H1: *Days with difficulties present* (out of 30 days) were significantly reduced after 8 weeks on Heylo™ compared to standard of care from 8.2 to 6.8 (LS mean difference -1.4, 95% CI: -2.7; -0.2, p=0.025) (Figure 12). Across a whole year, this corresponds to 17 days, equalling 3.4 work weeks with no difficulties present reported by subjects when using Heylo™.

H2: *Days totally unable to carry out activities or work because of any health condition* (out of 30 days) was non-significantly reduced after 8 weeks on Heylo™ compared to standard of care from 5.1 to 4.8 (LS mean difference -0.3, 95% CI: -1.4; 0.8, p=0.543) (Figure 12).

H3: *Days with cutting back or reducing usual activities or work because of any health condition* (out of 30 days) was non-significantly reduced after 8 weeks on Heylo™ compared to standard of care from 8.5 to 7.7 (LS mean difference -0.8, 95% CI: -2.1; 0.6, p=0.248).

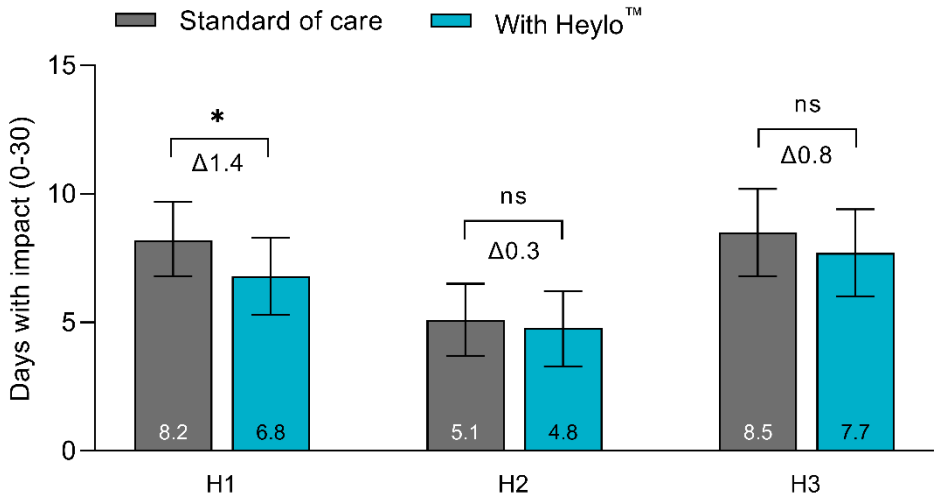


Figure 12 WHODAS 2.0 Days affected by disability.
LS mean values for days affected by disability with Heylo™ and standard of care are displayed (n=139). Error bars represent the 95% confidence intervals of the LS means. H1: Days with difficulties present (out of 30 days), H2: Days totally unable to carry out activities or work because of any health condition (out of 30 days), H3: Days with cutting back or reducing usual activities or work because of any health condition (out of 30 days). ∗P < 0.05, ∗∗P < 0.01, ∗∗∗P < 0.001, ns; not significant.

6.8.4 Feeling of security evaluated at the end of each test period.

The subjects were asked to rate the feeling of security while wearing the products on a 5-point scale.

The feeling of security increased significantly after 8 weeks when using Heylo™ compared to standard of care (p<0.001, n=133) (Figure 13). In all, 76% of the subjects had a good or very good feeling of security with Heylo™ and 58% without Heylo™ on standard of care. This corresponds to a 31% increase in subjects with a *good* or *very good* feeling of security with Heylo™.

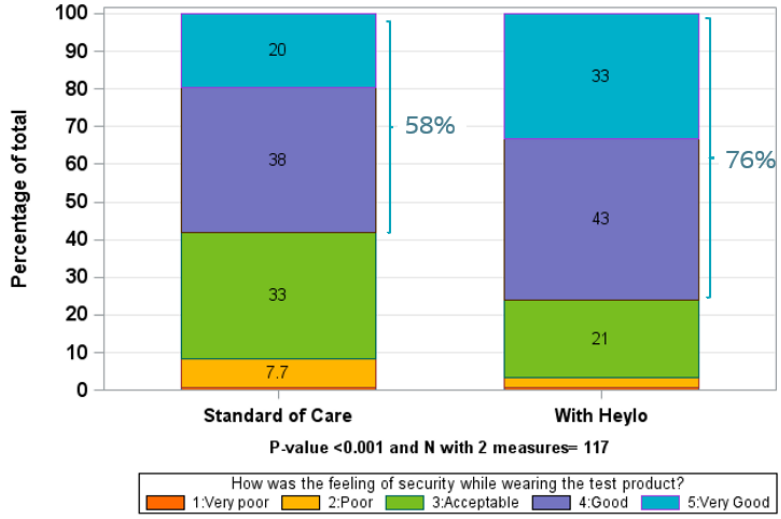


Figure 13 Feeling of security.
The proportion of subjects in the five categories is displayed for Heylo™ and standard of care (the figure only includes subjects with data on both time points, n=117, whereas the statistical analysis is performed on the full analyses set (with at least one data point, n=133).

6.8.5 Number of episodes of leakage outside the baseplate

A significant reduction in episodes of leakage outside the baseplate was observed after 8 weeks with Heylo™ when compared to standard of care (p<0.001) (Figure 14). The number of episodes of leakage outside the baseplate, after 8 weeks decreased from 2.3 leaks per 2 weeks with standard of care to 1.6 leaks per 2 weeks with Heylo™, corresponding to a 31% reduction (95%CI: 15; 44).

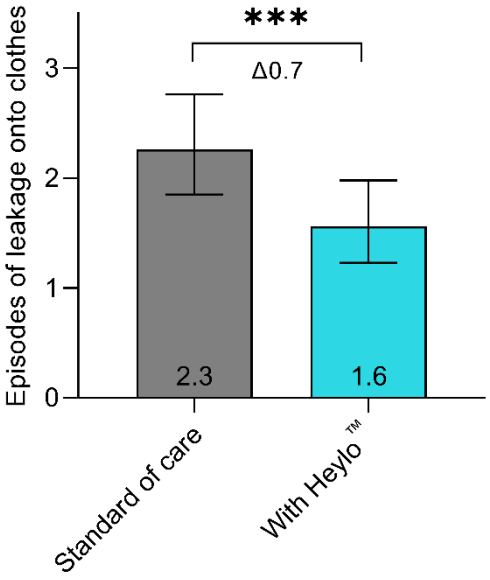


Figure 14 Number of episodes of leakage outside the baseplate.
LS mean values and a P-value for comparison of Heylo™ and standard of care are displayed (n=133). Error bars represent the 95% confidence intervals of the LS means. ∗P < 0.05, ∗∗P < 0.01, ∗∗∗P < 0.001.

6.9 Further evaluation of results and additional analyses of endpoints

6.9.1 Summary of efficacy metrics and effect sizes by Cohen's *d*

Table 10 summarises the efficacy analyses conducted as outlined in the statistical analysis plan, with the addition of standardized effect measures (by Cohen's *d*). The results of the standardized effect measures (Cohen's *d*) are presented in two ways. In the first column called 'Intra' the effect measures are divided by the intra subject SD. In the second column called 'Total' the effect measure is divided by the total population SD. Cohen's *d* calculated based on the intra-subject SD, is a measure of the effect a subject can expect from using Heylo™ instead of standard of care, relative to the standard deviation around the subject's general level on the scale. Cohen's *d* calculated based on the total SD is a population-based measure of the effect if the population uses Heylo instead of standard of care, relative to the standard deviation in the population. A commonly used interpretation of Cohen's *d* effect sizes is to refer to them as either small (*d*=0.2), medium (*d*=0.5), large (*d*=0.8) or very large (*d*=1.3), however these suggested values are arbitrary and should not be interpreted rigidly (37, 38). Larger effect sizes were observed for domains in the stoma-specific questionnaire (OLI tool) compared with the domains in the generic disability assessment (WHODAS 2.0).

Table 10 Summary of efficacy metrics

Results from the efficacy analyses after a test period of 8 weeks with Heylo™ or standard of care, provided as LS means with 95% CI, respectively, as well as LS mean differences and corresponding p-values. Moreover, standardized effect measures are provided as Cohen's *d* in two ways, (a) mean difference divided by within subject variability (SD) named '*Intra*', and (b) mean difference divided by total variability (group SD) named '*Total*'.

	After completion of 8 weeks test period				**Cohen's *d*** *(standardized effects)*	
	***with Heylo*™** LS mean (95% CI)	***with SoC*** LS mean (95% CI)	**LS Mean diff.** (95% CI)	**P-value**	Intra	Total
Ostomy Leak Impact tool						
Emotional Impact	73.4 (68.9; 77.9)	62.0 (57.6; 66.4)	11.42 (7.83; 15.01)	<0.001	0.79	0.44
Usual and Social Activities	81.1 (76.1; 86.1)	73.4 (68.6; 78.1)	7.71 (3.53; 11.90)	<0.001	0.59	0.32
Coping and in Control	73.6 (68.7; 78.6)	65.9 (61.0; 70.8)	7.77 (3.80; 11.74)	<0.001	0.49	0.27
WHODAS 2.0						
Domain 1: Cognition	21.7 (17.9; 25.4)	25.0 (21.3; 28.7)	-3.36 (-6.13; -0.59)	0.018	-0.31	-0.15
Domain 2: Mobility	26.0 (21.6; 30.5)	29.6 (25.2; 34.1)	-3.60 (-6.45; -0.74)	0.014	-0.32	-0.14
Domain 3: Self-care	17.2 (13.3; 21.0)	19.2 (15.4; 23.0)	-2.05 (-4.63; 0.54)	0.120	-0.20	-0.09
Domain 4: Getting along	27.2 (22.9; 31.6)	30.4 (26.1; 34.8)	-3.23 (-5.93; -0.52)	0.020	-0.30	-0.13
Domain 5: Life Activities	34.1 (29.1; 39.1)	38.9 (33.9; 43.8)	-4.79 (-8.33; -1.25)	0.008	-0.34	-0.17
Domain 6: Participation	33.0 (28.9; 37.0)	37.1 (33.1; 41.2)	-4.16 (-6.69; -1.63)	0.001	-0.42	-0.18
Total score	26.5 (22.9; 30.2)	30.0 (26.4; 33.6)	-3.49 (-5.49; -1.50)	<0.001	-0.44	-0.16
			Relative Risk (95% CI)			
Leakage onto clothes last 2 weeks	1.56 (1.23; 1.98)	2.26 (1.85; 2.76)	0.69 (0.56; 0.85)	<0.001		
			Proportional Odds ratio (95% CI)			
Feeling of security			0.427 (0.290; 0.629)	<0.0001		

6.9.2 Reasons and timing for subjects dropping out of the study

Reasons for participants dropping out of the study and timing of drop-outs are shown in Table 11. A total of n=8 participants dropped out of the study while using Heylo™ and n=7 participants dropped out of the study while being on standard of care. While using Heylo™, n=2 participants dropped out due to AEs related to the product, n=1 participant wished to discontinue, n=1 participant was lost to follow-up and n=4 participants dropped out due to AEs not related to the product. While using standard or care, n=3 participants dropped out of the study due to non-related AEs, n=3 participants wished to discontinue and n=1 was a protocol deviation (Table 11). All subjects dropping out of the study while using Heylo™ did so in the second test period. On average, subjects dropped out after 1.9 weeks [range 1 to 4 weeks] while on Heylo™ and after average 5.1 weeks [range 2 to 8 weeks] on standard of care.

Table 11 Reasons and time for drop-outs.

Drop-outs while on Heylo™ (n=8)			**Drop-outs while on standard of care (n=7)**		
Subject no	***Reason for Drop-out***	***Time of drop-out, weeks (W) & Test Period (TP)***	***Subject no***	***Reason for Drop-out***	***Time of drop-out, weeks (W) & Test Period (TP)***
020	AE (Skin Irritation, Causal relationship)	W10 (2 weeks in TP2)	016	AE (Peritonitis, Not related, Serious)	W2 (2 weeks in TP1)
068	AE (Surgery thyorid, not related, serious)	W11 (3 weeks in TP2)	055	Subject wishes to discontinue	V2 (W8) (at completion of TP1)
072	AE (contact dermatitis, Probably related)	V2 follow-up (W9) (1 week in TP2)	060	AE (Prolapse, Not related, Serious)	W6 (after 6 weeks of TP1)
075	AE (subject past away, Not related, Serious)	V2 follow-up (W9) (1 week in TP2)	063	Protocol deviation	W14 (after 6 weeks of TP2)
086	AE (Early stoma re-operation, Not related, Serious)	W12 (4 weeks in TP2)	087	AE (Morbus Crohn boost, Not related, Serious)	W10 (after 2 weeks of TP2)
104	Subject wishes to discontinue	W10 (2 weeks in TP2)	118	Subject wishes to discontinue	W4 (after 4 weeks of TP1)
135	AE (Died, Not related, Serious)	V2 follow-up call (W9) (one week in TP2)	129	Subject wishes to discontinue	W8 (at completion of TP1)
141	Lost to follow-up	V2 follow-up call (W9) (one week in TP2)			
Summary 2 related, not serious AEs 4 serious not related AEs 1 subject wishes to discontinue 1 lost to follow up All drop-outs are in the second test period after Soc in the first test period			**Summary** 2 not related, not serious AEs 1 serious, not related AE 3 subjects wishes to discontinue 1 protocol deviation 5 drop-outs in the first test period and 2 in the second test period		

6.9.3 Demographics and baseline values of subjects dropping out of the study and subjects completing the study

Table 12 shows demographics and baseline-values of endpoints of the drop-out subjects by Heylo™ and standard of care study groups, as well as for participants completing the study.
With the low number of subjects dropping out of the study being related to either use of Heylo™ or standard of care, it is not realistic to infer any relevant differences in baseline endpoint-values between drop-outs with Heylo™, drop-outs with standard of care and Completers of the study.

Table 12 Demographics and baseline values for endpoints for drop-outs (Heylo™ vs Standard of care) and Completers

Parameter	**Heylo™**	**Standard of Care**	**Completers of both test periods**
Demographics & competed study weeks	Drop-outs (n=8)	Drop-outs (n=7)	Completers (n=124)
Age in years (mean, SD, [range])	49.0, 18.7, [23–72]	53.1, 19.6, [18–81]	50.7, 12.8, [19–75]
Gender (n, % female)	6 (75%)	1 (14%)	71 (51.1]
Stoma-type (n, % ileo)	7 (88%)	3 (43%)	78 (63%)
Time with stoma (n, %, < 3 mths)	4 (50%)	3 (43%)	24 (19%)
Body mass index (kg/m^2) (mean, SD, [range])	23.7, 3.7, [19–29]	22.4, 4.7, [16–31]	26.2, 5.5, [15–44]
Work status (n, % sick leave)	1 (13%)	1 (14%)	20 (16%)
Total completed study weeks (w) when drop-out (mean, [range])	9.1 w, [9–12]	7.4 w, [2–14]	16 w, [16]
Completed study weeks (w) in the respective Test Period (mean, [range])	1.9 w, [1–4]	5.1 w, [2–8]	2 * 8 weeks
Baseline values for the endpoints			
Primary endpoint Emotional Impact domain of OLI (mean, SD, [range])	37.9, 16.0, [17–60]	65,2, 34.0, [10– 00]	51.8, 26.9, [0–100]
Secondary endpoint WHODAS 2.0 (participation domain) (mean, SD, [range])	40.1, 21.7, [8–71]	33.3, 30.0, [0–79]	39.5, 24.4, [0–96]
Exploratory endpoints			
Usual Social activity domain of OLI (mean, SD, [range])	61.4, 22.3, [25 89]	56.9, 23.7, [38–83]	62.4, 25.2, [0–100]
Coping and Control domain of OLI (mean, SD, [range])	38.5, 16.0, [17–58]	57.1, 35.8, [0–100]	57.4, 31.0, [0–100]
WHODAS 2.0 (Cognition) (mean, SD, [range])	11.3, 13.6, [0–40]	22.1, 24.0, [0–65]	20.4, 22.9, [0–90]
WHODAS 2.0 (Mobility) (mean, SD, [range])	25.0, 28.0, [0–75]	24.1, 22.1, [0–50]	25.5, 26.4, [0–100]
WHODAS 2.0 (Self-care) (mean, SD, [range])	16.3, 22.6, [0–50]	15.7, 21.5, [0–60]	16.1, 24.7, [0–100]
WHODAS 2.0 (Getting along) (mean, SD, [range])	18.8, 19.8, [0–58]	14.3, 20.8, [0–58]	25.4, 25.6, [0–100]
WHODAS 2.0 (Life activities) (mean, SD, [range])	39.4, 35.0, [0–92]	25.2, 29.7, [0–67]	35.6, 30.7, [0–100]
WHODAS 2.0 (Total score) (mean, SD, [range])	25.1, 21.2, [0-88]	22.5, 20.7, [0-58]	27.1, 21.0, [0-88]
Leakage episodes past 2 weeks (mean, SD, [range])	3.9, [1–10]	2.9, [0–14]	3.0, [0–28]
Feeling of security (n, %, with low or very low)	1 (13%)	1 (14%)	20 (16%)

6.9.4 Time trend evaluations

We subsequently investigated how some of the outcomes developed over time to have an impression of whether the effects seen between Heylo™ and standard of care (after 8 weeks) changed over time (week 4 or week 8). The evaluations are based on results from the pre-specified analyses of the endpoints.

6.9.4.1 Time trends for the three OLI domain scores

The differences in OLI scores between Heylo™ and standard of care for all three OLI domains: *Emotional impact*, *Usual and social activity* and *Coping and in control* were constant throughout the investigation and a significant change was observed after 2 weeks (p<0.001) (Appendix 12.16, Figure 1-3, respectively).

6.9.4.2 WHODAS 2.0 domain (Domain 1-6) scores and total WHODAS 2.0 score during the investigation

For domain 5 and the total WHODAS 2.0 score, a significant improvement was observed after 4 weeks with Heylo™ compared to standard of care (Appendix 12.16, Figure 8 and Figure 10). For domain 3 a significant improvement was seen after 4 weeks with Heylo™, however no significant improvement was noticed after 8 weeks (Appendix 12.13, Figure 6). For the remaining domains (1, 2, 4 and 6), significant improvements were first detected after 8 weeks with Heylo™ (Appendix 12.16, Figure 4-5, Figure 7 and Figure 9).

6.9.4.3 Episodes of leakage outside the baseplate during the investigation

The differences in episodes of leakage outside the baseplate were stable after 4 weeks of use (Appendix 12.16, Figure 11).

6.9.5 Severity range for WHODAS 2.0 Participation domain score

Lee et al. have divided the WHODAS 2.0 score (0-100) into disability severities (39). A mean of 30 on the WHODAS 2.0 disability score indicates that on average subjects are moderately impacted by their disability in daily life (39). A score from 0-4 meaning `No *problems*´, a score from 5-24 meaning `*Mild* disability´, a score from 25-49 meaning `M*oderate* disability´, and a score from 50-100 meaning `*Severe* disability´ (39)).

After 8 weeks 67% of the subjects were moderately or severely disabled without Heylo™ on standard of care and 56% of the subjects were *moderate* or *severe* disabled with Heylo™ (Figure 15). This corresponds to a 16% reduction in subjects with *moderate* or *severe* disability with Heylo™, measured by the WHODAS 2.0 *Participation* domain score.

The proportion of subjects with *no problems* or *mild* disability after 8 weeks was 33% without Heylo™ on standard of care and 44% with Heylo™ (Figure 15).

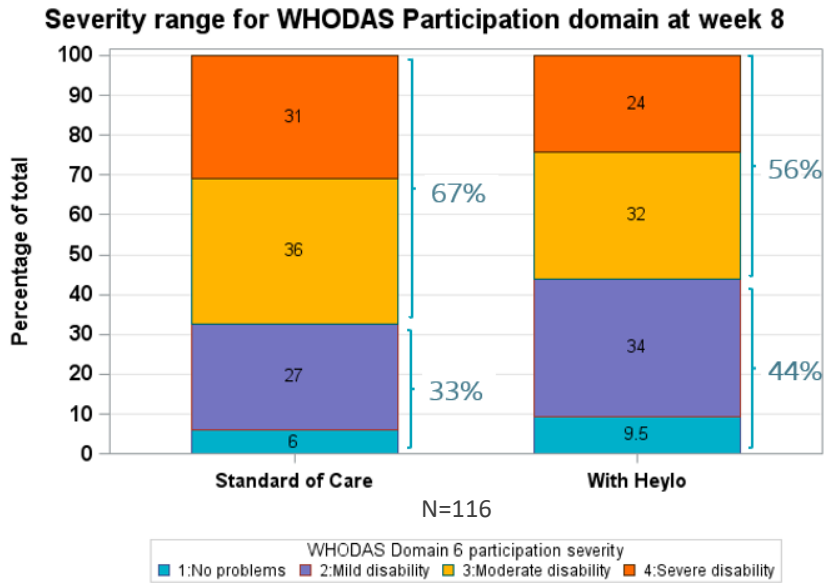


Figure 15 Severity range for WHODAS 2.0 *Participation* domain score.

Proportion of subjects in each of the different severity ranges for WHODAS 2.0 *Participation* domain score after 8 weeks with use of standard of care and Heylo™ are displayed (n=116, only including subjects with a measurement at both time points). Score from 0-4: No problems, score from 5-24: Mild disability, score from 25-49: Moderate disability, score from 50-100: Severe disability (39).

6.10 Post-hoc sensitivity analyses

6.10.1 Impact of ostomy product type, shape and time with a stoma

For the primary endpoint, the *Emotional impact* domain score, the sensitivity analysis showed a statistically significant and marginally higher effect of Heylo™ for 2-piece than 1-piece ostomy product users, and for convex users than for flat/concave users when compared to standard of care (data not shown, see Statistical Analysis Report - Tables - CP345 (40)). No significant difference was estimated for NPD vs Experienced users.

The sensitivity analyses for the secondary endpoint and leakage outside the baseplate showed similar effect sizes (not significantly different) for 2-piece and 1-piece ostomy solution users, for convex vs. flat/concave baseplate users and for NPD vs. Experienced users with Heylo™ compared to standard of care (data not shown, see Statistical Analysis Report - Tables - CP345 (40)).

6.10.2 Impact of time with a stoma evaluated in sub-group analyses

Post-hoc sensitivity subgroup analyses were performed for the three subgroups:

- NPDs
- Users with a temporary stoma
- NPDs with a temporary stoma

As described above, sensitivity analyses showed no significant difference between NPDs and Experienced users. The subgroup results were similar for the primary and the secondary endpoint as well as for the explorative endpoint leakage outside the baseplate when compared to the results based on the entire ITT population (data not shown, see Statistical Analysis Report - Tables - CP345 (40)). As an example, the estimated difference (LS mean difference) after 8 weeks for the primary endpoint was 11.4, 95% CI: 7.8; 15.0, p<0.001 for the entire ITT population (n=139) whereas it was 8.3, 95% CI: 1.3; 15.4, p=0.023, for the subgroup including NPDs (n=31), 9.1, 95% CI: 2.2; 15.9, p=0.009, for the subgroup including users with a temporary stoma (n=36) and 8.5, 95% CI: 0.6; 17.6, p=0.067 for the subgroup including NPDs with a temporary stoma (n=21). Many of the subjects are part of all three subgroups and therefore the results for the three subgroups are close to identical. This also applies to the secondary endpoint and for the explorative endpoint of numbers of leakages outside the baseplate (data not shown, see Statistical Analysis Report - Tables - CP345 (40)).

Impact of sex, age, type and shape of ostomy product and type of stoma evaluated in sub-group analyses

Post-hoc sensitivity subgroup analyses were performed for males and females separately. For the primary endpoint, the *Emotional impact* domain score, the score improved significantly for both males (LS mean difference 8.0, 95% CI: 3.6; 12.3, p<0.001, n=68) and females (LS mean difference 14.8, 95% CI: 9.3; 20.4, p<0.001, n=71). For the secondary endpoint, the *WHODAS 2.0 Participation* domain, the score tended to improve for males (LS mean difference -2.27, 95% CI: -5.89; 1.35, p=0.217, n=67) and significantly improved for females (LS mean difference -5.84, 95%CI: -9.43; -2.24, p=0.002, n=71).

Subgroup analyses for the primary endpoint was also performed separately for 1-piece and 2-piece users, for convex users and the group of flat and concave users as well as for subjects with ileostomy and colostomy. The *Emotional impact* domain score improved significantly for both 1-piece and 2-piece users (LS mean difference 7.6, 95% CI: 2.4; 12.8, p=0.004, n=61 and 14.1, 95% CI: 9.2; 19.0, p<0.001, n=78, respectively). The same was true for convex users and flat and concave users (LS mean difference 14.2, 95% CI: 9.2; 19.3, p<0.001, n=88 and 7.1, 95% CI: 2.7; 11.6, p=0.002, n=51, respectively) as well as for subjects with ileostomy and for subjects with colostomy (LS mean difference 10.6, 95% CI: 5.8; 15.4, p<0.001, n=88 and 12.9, 95% CI: 7.6; 18.2, p<0.001, n=51).

The subgroup analysis of the *Emotional impact* domain score for subjects older than 65 years only included 16 subjects. Still an improvement was estimated with a LS mean difference of 8.3, 95% CI: -0.3; 17.0, even if

it was not statistically significant (p=0.06). In the subgroup including subjects ≤ 65 years the score improved significantly (LS mean difference 11.8, 95% CI: 7.9; 15.7, p<0.001, n=68) (Statistical Analysis Report – Tables – CP345 (40)). For the secondary endpoint the *WHODAS 2.0 Participation* domain the score improved significantly for those below 65 years (LS mean difference -5.26, 95% CI:-7.94; -2.59, p<0.001, (n=122)), but not for those older than 65 years (LS mean difference 4.18, 95% CI: -3.36; 11.71, p=0.268, (n=16)), however the power of the test is low due to the low number of subjects.

6.10.3 Study effect (impact of test period)

For most of the endpoints, a significant effect of test period was estimated in the pre-specified analyses showing better results in period 2 no matter what product (Heylo™ or standard of care) the subject started on. The estimated effect of period, which might be due to a study effect over time, is not expected to affect the estimated difference between Heylo™ and standard of care because of the randomised cross-over study design.

6.10.4 Responder analyses for the primary and secondary endpoints

Two types of responder analyses were performed for the primary and secondary endpoints:
(a) Responder analyses with absolute thresholds and
(b) Responder summary based on changes from baseline (improving, deteriorating or no change).

(a) Responder analyses with absolute thresholds
For the primary endpoint *Emotional Impact* domain score, the responder analysis, using a responder success threshold of a score ≥67, showed that the estimated marginal proportions of responder success were 68% for the Heylo™ group and 44% for the standard of care group after 8 weeks. The corresponding odds ratio is estimated to 2.68, 95%CI:1.59;4.51, p<0.001 (Table 13).

For the secondary endpoint, the *WHODAS 2.0 Participation* domain score, the responder analysis, using a responder success threshold of a score <25, showed that the estimated marginal proportions of responder success were 44% for the Heylo™ group and 30% for the standard of care group after 8 weeks. The corresponding odds ratio is estimated to 1.79, 95%CI:1.05;3.05, p<0.035 (Table 13).

For both endpoints, the responder analyses showed that after 8 weeks of usage there was a statistically significantly difference between Heylo™ and standard of care groups, favouring Heylo™.

Table 13 Responder analysis with absolute thresholds.

A responder was for the Emotional Impact domain defined as a subject with a score ≥67, which corresponds to the mean score for the background stoma population. For the WHODAS participation domain, a subject was considered a responder if the score was <25 (either 0-4 (No problems), or 5-24 (mild disability)).

	Responders* after completion of 8 weeks test period			
	With Heylo™ Proportion (%)	*With SoC Proportion* (%)	**Odds Ratio** (95% CI)	**P-value**
Ostomy Leak Impact tool Emotional Impact	68%	44%	2.68 (1.59;4.51)	p<0.001
WHODAS 2.0 Participation (domain 6)	44%	30%	1.79, (1.05;3.05)	p<0.035

(b) Responder summary based on changes from baseline

The proportion of responders (any increase from the baseline score) for the *Emotional Impact* score while on Heylo™ was around 80% after just 2 weeks of product usage (Table 14). This level of responders continued throughout the 8 weeks test period. The proportion that worsened while on Heylo™ was between 10-19% across the 8 weeks test period. In comparison, the proportion of responders while on standard of care was 62% after 2 weeks and stayed between 64%-67% during the 8-week test period. The proportion that worsened from baseline while on standard of care was around 28% throughout the test period. These proportions are visually presented in corresponding Sankey diagrams, where also the change between groups over time are indicated (Figure 16).

The proportion of responders (any improvement from the baseline score) for the WHODAS 2.0 *Participation* (domain 6) while on Heylo™ increased from 51% at week 4 to 55% at week 8 (Table 15). The proportion that worsened while on Heylo™ decreased from 38% at week 4 to 30% at week 8. In comparison, the proportion of responders while on standard of care was 49-50%. The proportion that worsened while on standard of care was around 40-41% throughout the test period. These proportions are visually presented in corresponding Sankey diagrams (Figure 17).

Table 14 Responder summary based on changes from baseline for Emotional Impact domain

Proportion of responders (subjects improving, deteriorating, or stayed unchanged compared with baseline) across study timepoints for the Emotional Impact score while on Heylo™ and standard of care, respectively.

	Time in test period			
	2 weeks n (%)	4 weeks n (%)	6 weeks n (%)	8 weeks n (%)
Change while on *Heylo*™				
Total	n=123	n=119	n=120	n=122
Improving from Baseline	98 (79.7)	90 (75.6)	100 (83.3)	99 (81.1)
No change from Baseline	9 (7.3)	7 (5.9)	7 (5.8)	11 (9.0)
Worsening from Baseline	16 (13.0)	22 (18.5)	13 (10.8)	12 (9.8)
Change while on *Standard of Care*				
Total	n=134	n=131	n=126	n=128
Improving from Baseline	83 (61.9)	88 (67.2)	81 (64.3)	83 (64.8)
No change from Baseline	14 (10.4)	6 (4.6)	10 (7.9)	8 (6.3)
Worsening from Baseline	37 (27.6)	37 (28.2)	35 (27.8)	37 (28.9)

Table 15 Responder summary based on changes from baseline for Participation domain

Proportion of responders (subjects improving, deteriorating, or stayed unchanged compared with baseline) for the WHODAS 2.0 *Participation in society* (domain 6) while on Heylo™ and standard of care, respectively.

	Time in test period			
	2 weeks n (%)	4 weeks n (%)	6 weeks n (%)	8 weeks n (%)
Change while on *Heylo*™				
Total	N/A*	n=112	N/A*	n=121
Improving from Baseline	N/A*	57 (50.9)	N/A*	67 (55.4)
No change from Baseline	N/A*	12 (10.7)	N/A*	18 (14.9)
Worsening from Baseline	N/A*	43 (38.4)	N/A*	36 (29.8)
Change while on *Standard of Care*				
Total	N/A*	n=119	N/A*	n=128
Improving from Baseline	N/A*	58 (48.7)	N/A*	64 (50.0)
No change from Baseline	N/A*	13 (10.9)	N/A*	12 (9.4)
Worsening from Baseline	N/A*	48 (40.3)	N/A*	52 (40.6)

N/A: Not available. *WHODAS 2.0 was collected every fourth week.

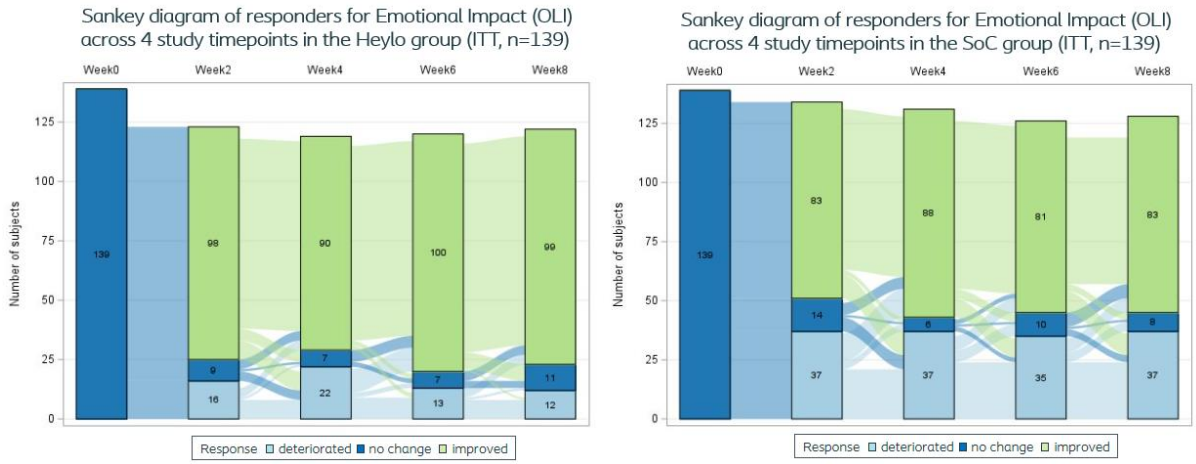


Figure 16 Sankey diagram for Emotional Impact

The Sankey diagrams show the number of subjects with a change in the Emotional Impact (OLI) score from baseline across four timepoints. The left figure shows responders while testing Heylo™ for 8 weeks. The right figure shows responders while on standard of care for 8 weeks. Green represents subjects who improved from their baseline value. Dark blue represents subjects with unchanged scores. Light blue represents subjects who scored lower (worsened) from the baseline (exact value). The dark blue bar (at week 0) indicates the number of patients with a baseline measure, and the light blue bar (between week 0 and 2) indicates the number of subjects with a measure after 2 weeks.

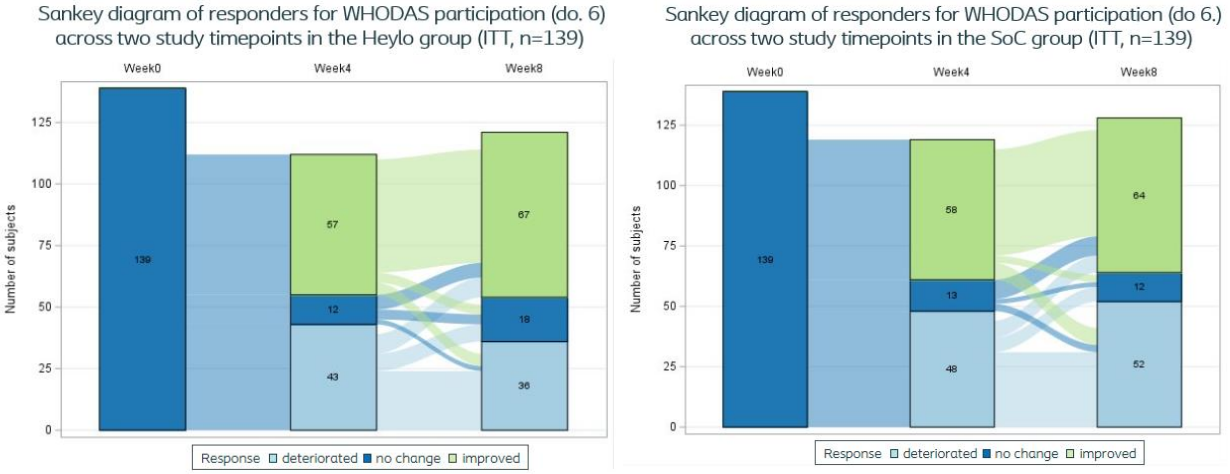


Figure 17 Sankey diagram for *Participation* domain

The Sankey diagrams show the number of subjects with a change in the WHODAS 2.0 *participation* (domain 6) score from baseline across two timepoints. The left figure shows responders while testing Heylo™ for 8 weeks. The right figure shows responders while on standard of care for 8 weeks. Green represents subjects who improved from their baseline value. Dark blue represents subjects with unchanged scores. Light blue represents subjects who worsened from the baseline (exact value). The dark blue bar (at week 0) indicates the patient number with a baseline measure, and the light blue bar (between week 0 and 4) indicates the number of subjects with a measure after 4 weeks.

6.10.5 Analyses of the primary and secondary endpoint based on subjects with observations in both periods

The assumption for imputation of missing values in the statistical analyses is that missing data are missing completely at random. The effect of the imputation done by the linear mixed model for the primary and secondary endpoints were tested in additional analyses that only included subjects who had measurements in both test periods at Week 8 (n=117 and n=116, respectively). The results of these analyses were minimally different from the results of the pre-defined analyses (primary endpoint: LS mean difference 10.9, 95% CI: 7.2; 14.7, p<0.001 (n=117); secondary endpoint: LS mean difference -3.8, 95% CI: -6.4; -1.2, p=0.004 (n=116)). This indicates that the imputation of missing values in the primary analyses did not affect the results markedly.

6.10.6 Analysis of all randomized subjects with imputation of missing data

The estimated week 8 treatment contrasts obtained when using MI assuming MAR (Table 16), were similar to the estimated week 8 treatment contrasts from the pre-specified repeated mixed measure model analyses (Table 10). When reference-based imputation was applied, the LS mean with Heylo™ was a little closer to the LS mean with standard of care, and the estimated treatment effect was hence diminished. In all the performed analyses the treatment effect was significantly different from 0 based on the pre-specified significance level of 2.5%. Inclusion of baseline measures to the models had little impact on the estimated treatment contrasts. Inclusion of baseline data impacted Cohen's *d* values based on the total variance, as the baseline explains some of the variance between subjects.

Table 16 Efficacy estimates with imputation of missing data

Efficacy estimates from the Linear Mixed Model (LMM) analyses of week 8 data from all randomized subjects based on multiple imputation of missing data. The two imputation methods described in section 5.3.20.7.6 have been applied.

All randomized subjects (n=144) with imputed data	After completion of 8 weeks test period				Cohen's *d* (standardized effects)	
	with Heylo™ LS mean (95% CI)	***with SoC*** LS mean (95% CI)	**LS Mean diff.** (95% CI)	**P-value**	Intra[#]	Total[¤]
Ostomy Leak Impact (OLI) Emotional Impact Score						
MI Reference based	73.2 (68.7; 77.7)	62.7 (58.2; 67.1)	10.55 (6.22; 14.87)	<0.001	0.60	0.39
MI Reference based baseline inc.	73.2 (69.5; 76.9)	62.7 (59.1; 66.3)	10.53 (6.18; 14.87)	<0.001	0.59	0.49
MI assuming MAR	74.2 (69.7; 78.7)	62.7 (58.3; 67.1)	11.49 (7.17; 15.81)	<0.001	0.65	0.43
MI assuming MAR baseline inc.	74.2 (70.5; 77.9)	62.7 (59.1; 66.3)	11.51 (7.16; 15.85)	<0.001	0.65	0.54
WHODAS 2.0 Participation (domain 6)						
MI Reference based	33.1 (29.0; 37.2)	36.9 (32.9; 41.0)	-3.83 (-6.99; -0.67)	0.018	-0.30	-0.16
MI Reference based baseline inc.	33.2 (29.9; 36.4)	36.9 (33.7; 40.1)	-3.76 (-6.92; -0.59)	0.020	-0.30	-0.20
MI assuming MAR	32.7 (28.6; 36.8)	37.0 (32.9; 41.0)	-4.25 (-7.51; -0.99)	0.011	-0.34	-0.18
MI assuming MAR baseline inc.	32.8 (29.5; 36.0)	36.9 (33.8; 40.1)	-4.14 (-7.40; -0.88)	0.013	-0.33	-0.22

[#] Mean difference divided by within ‚Intra' subject SD; [¤] Mean difference divided by Total SD. In the models with baseline included, baseline as well as the interaction between baseline and period were included. Inc.=Included

6.11 Safety assessments

A total of 20 AEs were documented in 18 subjects (18/144*100 = 12.5%), including 11 non-related serious AEs (11/144*100 = 7.6%) and 9 non-serious AEs (9/144*100 = 6.3%).

AEs that were classified as "causal relationship", "probably related" or possibly related to a device were treated as "related". Of the 9 non-serious AEs, 5 AEs in 5 different subjects (5/144*100 = 3.5%) were "related" to Heylo™ including two incidents of potential contact dermatitis and three skin irritations. See Table 18 for further description.

Table 17 lists AEs (ADE) and serious AEs (SADE) "related" to Heylo™ and standard of care divided into mild, moderate, severe, or not specified.

Table 17 Intensity of AE's and SAEs "related" to the investigational and non-investigational devices.

	Heylo™		Standard of care	
Intensity	**ADE**	**SADE**	**ADE**	**SADE**
Mild	0	0	0	0
Moderate	4	0	0	0
Severe	1	0	0	0
Not specified	0	0	0	0
Total	5	0	0	0

Source: Summary from EOT

Table 18 describes all the device-related AEs.

Table 18 Description of ADEs.

ADE description	*Heylo*™ (subject ID)	Start date	Out-come date	Comment (treatment, confounding factors)	Outcome
Skin Irritation	020	05 APR 2022	19 APR 2022	Adhesive surface of the sensor layer has made skin irritations	Resolved
Contact dermatitis	072	23 JUN 2022	-	Skin irritation under the sensor layer	Ongoing
Skin Irritation under the layer. May be allergy.	078	10 AUG 2022	26 SEP 2022	Subject used a cream that relieved the skin irritation. After using Heylo significant improvement.	Resolved
Skin irritation	101	16 JUN 2022	22 JUN 2022	Skin irritation under the sensor layer around the stoma	Resolved
Contact dermatitis	133	17 JUL 2022		Contact dermatitis on the sensor layer (layer on the transmitter, outside of the baseplate), skin reactions present.	Ongoing
Total	**5**				

Source: Summary from EOT

The investigator was responsible for classifying each AE into serious or non-serious and causal relationship with the investigational device or procedure. The classification was agreed upon by the sponsor.

Principal investigator ensured that adequate medical care was provided to subjects experiencing an AE during or after participation in the clinical investigation. All serious AEs were followed until a resolution was addressed.

Further details regarding AEs are available in the Statistical Analysis Report - Listings - CP345 (34).

6.11.1 Device deficiencies

In all, 2 device deficiencies on Heylo™, distributed across 2 subjects were observed in this investigation. Table 19 shows the observed device deficiencies. None of the reported 2 device deficiencies could have led to a SADE. Consequently, no corrective actions were taken.

Table 19 Device deficiencies

Observed device deficiency	Action
Transmitter defect	*NA*
Transmitter defect	The subject used the replacement transmitter.

7 Discussion and overall conclusions

Leakage is a serious issue amongst people living with a stoma (1, 41) and early detection may be a simple mean of addressing this partly inevitable complication. This clinical trial was designed to study the positive care effect of a novel app-driven digital leakage notification system (Heylo™) on the QoL and disease burden of ostomy carriers characterized using the *Emotional impact* domain and the *Participation* domain from the previously validated OLI and WHODAS 2.0 questionnaires respectively (16-18). To achieve our goal, a randomized-controlled cross-over study comparing Heylo™ to standard of care after 8 weeks of product usage was performed. The use of Heylo™ was associated with a statistically significant improvement in subjects QoL and disease burden after 8 weeks in comparison to standard of care. More specifically, while using Heylo™ subjects overall reported less embarrassment, better sleep, and less worry about leakage, and they reported they were better able to live with dignity, be with friends and family, and participate in social life activities.

7.1 Strengths and limitations

The results of this clinical investigation should be interpreted considering some strengths and limitations of the study design.

A major strength of this investigation is the randomized cross-over study design. The randomization of subjects to the two treatment sequences limited variations and thus potential influence from confounding factors. Moreover, the cross-over design allowed subjects to serve as their own control, thereby allowing us to disregard the between subject variability. Also, no carry-over effect was expected in the investigation, as the duration of the investigation was 2 times 8 weeks, and the main evaluation was done by the end of each test period.

The investigation was conducted in a real-world setting as the homecare nurses performed all study visits at the subjects' homes or remotely by a tele-health call. Moreover, subjects were not instructed on when to change their ostomy product, which means that they could follow their normal change routine or change due to a notification from the app, if relevant. Therefore, the 'actual' observed care effect of Heylo™ is close to what can be expected in real life.

Another strength of the present investigation was that the number of subjects who dropped out was lower than expected and was equally distributed between Heylo™ and standard of care (n=8 while on Heylo™ and n=7 while on standard of care).

In addition, the inclusion of multiple sites (homecare nurses) distributed across Germany strengthens the generalisability of the findings.

The clinical data of the study participants regarding the underlying pathology as indication for ostomy surgery is largely in accordance with the current literature (42). Therefore, the study population was a true reflection of the real-life scenario. This is also true with respect to the demographic characteristics of the study population.

All patients completed the questionnaires individually on their smart phones every 2nd week without external aid and independently from the nurse visits and follow-up calls. Thus, questionnaires have been filled in consistently across the study periods and therefore nurses are not expected to have had any direct influence on the subjects' responses to the questionaries. Thus, the outcomes reported in this investigation are truly patient reported. Still, a remaining limitation exists in the non-blinded study design. The fact that both the subjects and the nurses were not blinded could potentially, consciously or subconsciously, have affected the subjects' responses to the questionnaires.

In addition, study participants were recruited from a prospectively maintained Coloplast care database and thereby most subjects used an ostomy solution from Coloplast as their current solution (standard of care). Thus, the proportion of subjects using Coloplast appliances is higher in the present investigation than for the general stoma population. However, it has previously been shown (CP340) that Heylo™ works as well with

baseplates from other brands, and thereby the results generated in this study is not expected to be influenced in any direction (11).

Nonetheless, results of this investigation suggest that Heylo is an effective leakage notification system with a huge potential of improving the QoL of stoma carriers.

7.2 Demographics

The subjects in this investigation were evenly distributed between males and females. The age of the subjects ranged from 18 to 81 years and the average age was 50.7 years. The average age of a subject in this investigation is slightly younger than in the previous investigation CP321 (10) and in other clinical investigations with stoma care products (43-45), however, is in line with the investigation CP340 (11). This probably reflects that an inclusion criterion for participation in this investigation was that subjects should have a smartphone compatible with the Heylo™ app and that the investigation included both new stoma users, who had had their stomas for less than 3 months and experienced users who had had their stomas for more than 3 months. On average the subjects had had their stoma for 6.1 years.

Another inclusion criterion of the investigation was that the subject should at least to some degree worry about leakage. In the general stoma population, 67% of the individuals living with a stoma worry to '*some or higher degree*' about leakage (2, 3). The general leakage worry among people with a stoma corroborates to that leakage and leakage worry are common problems for people with a stoma (2, 3), suggesting that the study participants of this investigation were representative of many end-users.

Collectively, the generalisability of the population enrolled in this clinical investigation was considered good as it overall represented the general population of people living with a stoma.

7.3 Primary endpoint

The primary endpoint of this investigation was the *Emotional Impact* domain score measured by the OLI tool (section 5.3.8.1).

A significant improvement in mean *Emotional impact* score was found after 8 weeks with Heylo™ compared to standard of care. The score increased by 11.4 points with Heylo™ compared to standard of care.

The observed improvement is considered clinically relevant as it is higher than the previously described MCID of Δ7.6 (mean of three MCIDs described from three different methods) for this domain (16). This indicates that Heylo™ provides a meaningful change for subjects in how much their QoL is impacted by leakage and that the subjects felt less frustration, less embarrassment, less panic and had better sleep with Heylo™. Furthermore, standardized effect measures, reported as Cohen's *d*, indicate medium to large positive effect of Heylo™ on the *Emotional impact* domain, when comparing to stated thresholds in the literature (37, 38).

In addition, the significant difference in *Emotional impact* score between Heylo™ and standard of care was constant during the investigation, evident at 2 weeks. These results indicate that people living with a stoma quickly will benefit from using Heylo™, as the observed positive effect of Heylo™ was seen early after study initiation.

The impact on the *Emotional impact* domain score when using Heylo™ has also previously been investigated in two clinical studies conducted in Denmark (CP321, a 3-week evaluation) and the UK (CP340, a 12-week evaluation). In both studies, the use of Heylo™ lead to significant improvements in the *Emotional impact* domain score, highlighting that the effect of Heylo™ is apparent across countries and with different clinical setups. In the UK investigation (CP340), a significant and clinically relevant difference of Heylo™ on the *Emotional impact* domain score was also observed after short-term use, supporting the outcome of the present investigation that users quickly experience improvements in QoL once using the product.

7.4 Secondary endpoint

The secondary endpoint was the *Participation* domain score measured by WHODAS 2.0 (section 5.3.8.2).

A significant improvement in mean *Participation* domain score was found after 8 weeks with Heylo™ compared to standard of care. The score decreased by 4.2 points. This result indicates an overall better participation in society with Heylo™, including an ability to obtain personal dignity.

Even though WHODAS 2.0 (36-item) was developed more than a decade ago (17, 18), a single MCID score for the WHODAS 2.0 (36-item) has not yet been established (20). MCID-values are sensitive to different populations and clinical scenarios, thus a range of MCID-value estimates exist for a given domain measure depending on the context for which it is used in (23-25).

To our knowledge, only few studies have established MCID-values for the individual WHODAS 2.0 (36-item) domain scores and the total score, in populations and clinical scenarios comparable to the present study (21, 22).

In two studies of patients with lower back pain and hip and knee osteoarthritis (using the Polish version of the WHODAS 2.0 (36-item), the MCID values for the *Participation* domain were Δ2.55 (22) and Δ4.62 (21), respectively.

The observed improvement in the WHODAS 2.0 *Participation* domain score was in our investigation Δ4.2, which is higher than the average MCID-value of the two above-mentioned studies (Δ3.59) (21, 22). This indicates that Heylo™ provides a meaningful change for subjects in this domain. Furthermore, standardized effect measures, reported as Cohen's *d*, indicate small to medium positive effect of Heylo™ on the *Participation* domain (WHODAS 2.0), when comparing to stated thresholds in the literature (37, 38).

The mean WHODAS 2.0 Participation score of the general population across 19 countries is approximately 6. This indicates that people with a stoma generally show difficulties in participating in society compared with the general population (17, 18).

As descried in the previous section 6.9.5, Lee et al. have divided the WHODAS 2.0 score (0-100) into four disability severities (39). Scores from 0-4 reflect `*No problems*´, scores from 5-24 indicate `*Mild* disability´, scores from 25-49 reflect `M*oderate* disability´, whereas scores from 50-100 imply that subjects are `*Severely* disabled´ in their daily life (39). A mean of 30 on the WHODAS 2.0 disability score indicates that on average subjects are medium impacted by their disability in daily life (39).

In this investigation, we observed that after 8 weeks, 67% of the subjects were moderately or severely disabled without Heylo™ (on standard of care) and 56% of the subjects were *moderately* or *severely* disabled with Heylo™. This corresponds to a 16 percent reduction in number of subjects with *moderate* or *severe* disabilities with Heylo™. The average disability value of the WHODAS 2.0 *Participation* domain score in our investigation was comparable with the value of the WHODAS 2.0 *Participation* domain score (12-item version) observed in the study by Lee et al., assessing disability for colorectal cancer survivors with or without a stoma (39). In the study by Lee et al. colorectal cancer survivors living with a stoma scored significantly worse in the *Participation* domain (32.3 vs. 22.2) and in the *Mobility* domain (31.1 vs. 20.3) compared with the group of colorectal survivors without a stoma, while no difference in the total score was observed between the two groups. These results highlight that people living with a stoma are generally quite disabled in the *participation* domain, which means that they are less able to participate in society and that they feel less able to live a life with dignity compared to people without a stoma. Nonetheless, the use of Heylo™ showed a reduction in the number of subjects being moderate or severely disabled in participating in society in everyday life, which was also highlighted in the accompanying responder analysis favoring the use of Heylo™.

7.5 Exploratory endpoints

7.5.1 OLI domains *Usual and social activities* and *Coping and in control*

As previously mentioned, the questions in the OLI tool are grouped into three domains which concerns the *Emotional impact* of having a stoma (primary endpoint), *Usual and Social activities*, as well as *Coping and in control*.

Like the results of the OLI domain score for *Emotional impact* (the primary endpoint), the mean score for the two other OLI domains *Usual and social activities* and *Coping and in control* also improved significantly after 8 weeks with Heylo™ compared to standard of care.

The score for the domain *Usual and social activity* increased by 7.7 points and the score for the domain *Coping and in control* increased by 7.8 points. The observed improvements in these domains were higher than the described relevant MCIDs of Δ6.6 and Δ7.2, respectively (when using the mean of three MCIDs through use of three different standard methods) (16). These results add to the relevance of Heylo™ to provide a meaningful change for subjects in how much their QoL is impacted by leakage and that the subjects overall felt less embarrassment, less frustration, better engagement in social activities and felt better able to cope with and control their situation.

Moreover, the changes to these two OLI domain scores when using Heylo™ were significant just after 2 weeks of intervention and stayed stable throughout the intervention.

The findings, of the present investigation, of improvements across all three OLI domains are in line with the results of the recently conducted investigation CP340 (11). In the CP340 investigation, patients who recently had their stoma formed (stoma age < 9 months) also showed improvements in all three domains of the OLI and again with the largest effect size recorded for the Emotional impact domain.

7.5.2 WHODAS 2.0 domain scores and total WHODAS 2.0 score

As previously mentioned, the WHODAS 2.0 disability tool is based on 36-item questions covering 6 different disability domains.

For each domain, the domain score ranges from 0-100. All six domain scores can also be summarized, into a total WHODAS 2.0 score to measure the change in overall disability.

Five out of the six WHODAS 2.0 domain scores and the total WHODAS 2.0 score improved significantly after 8 weeks with Heylo™, compared to standard of care, reflected as a decrease in the average scores.

The self-care domain score did not decrease significantly after 8 weeks, although a small tendency towards a decrease was observed. However, since the use of Heylo™ is also only expected to have a positive impact on one or maybe two of the item questions in this domain (those related to hygiene), a significant change in this domain was also not expected considering the short intervention. Interestingly, the average score of 19.2 for this domain indicates that subjects, in general, were impacted in their ability to self-care, compared to the general population, which may be due to underlying conditions, age, comorbidities, or due to having a stoma.

The effect sizes recorded in the present investigation for the WHODAS 2.0 domains were generally in the range of the MCID-values previously established (see discussion in *Secondary Endpoint*), indicating that Heylo™ provides a meaningful change across multiple domains for subjects (Table 20).

Table 20 Effects sizes and MCID-values

	Effect sizes CP345	MCID		
WHODAS Domains 1-6		Low back pain (22)	Osteoarthritis (21)	Average
Domain 1: Cognition	3.4	1.71	0.89	1.30
Domain 2: Mobility	3.6	7.93	5.15	6.54
Domain 3: Self-care	2.1	5.67	2.55	4.11
Domain 4: Getting along	3.2	4.85	2.52	3.69
Domain 5: Life activities	4.8	6.07	2.83	4.45
Domain 6: Participation	4.2	4.62	2.55	3.59
Total WHODAS 2.0 score	3.5	4.87	3.29	4.08

The total WHODAS 2.0 score decreased by 3.5 points after 8 weeks with Heylo™ compared to standard of care. This observed improvement was within the range of previously reported MCIDs for the total WHODAS 2.0 of 3 - 5 points, as described in the literature (21, 22, 25, 26), and strengthens the relevance of Heylo™ to provide a meaningful change for subjects.

In addition, the observed change in WHODAS 2.0 domain scores was for some of the domains (3, 5 and the total domain score) already present after 4 weeks with Heylo™.

In the present investigation, a positive care effect of Heylo™ after 8 weeks was demonstrated with both the disease-specific tool (OLI) and the generic WHODAS 2.0 tool measuring differences in disability degrees. A similar finding was observed in the previously conducted investigation CP340, where clinical benefits of Heylo™ (CP340) were also demonstrated with the use of both the disease-specific OLI tool and a generic health-related QoL tool, EQ-5D-5L/EQ-VAS (11).

At the end of the WHODAS 2.0 questionnaire subjects were additionally asked three questions to estimate how many days they during the past 30 days had had difficulties in disability to various degrees. A significant reduction in the estimated number of days (1.4 days out of 30 days) with difficulties present was reported by subjects after 8 weeks when using Heylo™, corresponding to approximately 17 days in a year, which almost equals three and a half work weeks. No statistical differences were observed between Heylo™ and standard of care in reported days where the subjects were completely unable to carry out activities or work because of any health condition (out of 30 days) nor in days with cutting back or reducing usual activities or work (out of 30 days). Even though our investigation did not suggest that people experienced differences in number of days with reduced ability to work or carry out activities when using Heylo™ compared with standard of care, the finding of a significant reduction in number of days with difficulties present is important, and it corroborates previous findings, where it was reported that 15% are highly impacted in their ability to work or completely prevented subjects from working due to leakage of stomal effluent and the worry hereof (3).

7.5.3 Feeling of security evaluated at the end of each test period.

The feeling of security increased significantly after 8 weeks when using Heylo™ compared to standard of care. In all, 76% of the subjects had a *good* or *very good* feeling of security with Heylo™ vs. 58% on standard of care. This corresponds to a 31% increase in subjects with a *good* or *very good* feeling of security with Heylo™. That people experienced an increased feeling of security with Heylo™ corroborates our observations in the previous clinical investigations of the product (CP321 and CP340) (10, 11).

A recent study on an international stoma population reported a positive correlation between worry about leakage and the use of specific stoma care accessories (i.e. rings/seals, paste, tape and belts); the more individuals worried about leakage, the more accessories they used (3).

7.5.4 Episodes of leakage incidence outside the baseplate

A significant reduction in episodes of leakage outside the baseplate was observed after 8 weeks with Heylo™ compared to standard of care. The number of episodes of leakage incidence outside the baseplate, after 8 weeks decreased from 2.3 leaks per 2 weeks with standard of care to 1.6 leaks per 2 weeks with Heylo™, corresponding to a 31% reduction.

Also, the differences in episodes of leakage outside the baseplate was stable after 4 weeks of use (see appendix 12.16, Figure 11).

The result of this investigation is in line with the results of leakage episodes of previously conducted clinical studies (CP321 and CP340) which both demonstrated that subjects experienced significantly fewer incidents of leakage outside the baseplate when using Heylo™ compared to the period preceding study initiation (10, 11). The significant reduction in the patient reported rate of leakage episodes with Heylo™ is very much associated with the efficacy of the system to detect a threatening leakage, thereby giving the ostomy carrier enough time to act before leakage becomes evident.

Multiple studies have highlighted that the frequency with which subjects experience leakage outside the baseplate is associated with QoL (3-5) and disutility (6). Leakage of stomal effluent progressing outside the baseplate (e.g. onto clothes or bed sheets) can be socially embarrassing, is often distressing and is always inconvenient for the individual to experience, with more than 90% of people living with a stoma worrying about leakage (2, 3). Subjects use different means to mitigate the risk of experiencing future leakage incidents, such as increasing the use of ostomy solutions (bag, baseplate and supporting products), and some have consultations with health professionals (7). Implementation of stoma care innovations that can reduce the number of leakage incidents outside the baseplate and related concerns may potentially limit the overuse of other stoma care products people use to mitigate the risk of leakage.

7.5.5 Post-hoc analyses

Post-hoc analyses for the primary endpoint and secondary endpoint were conducted with all randomized subjects using two imputation methods for missing data. The post-hoc analyses with all randomised subjects confirmed the conclusions from the pre-specified analyses of the primary- and secondary endpoints, by showing significant differences between Heylo™ and standard of care after 8 weeks for both endpoints.

Responder-analyses with absolute thresholds for the primary- and secondary endpoint showed significant higher proportions of subjects with responder-success when using Heylo™ compared with standard of care after 8 weeks, thus supporting the conclusions from the pre-specified analyses of the primary- and secondary endpoints.

7.6 Safety evaluation

In all, 5 AEs in 5 different subjects (5/144 *100 = 3.5%) were related to the investigational device, however, none of these were classified as serious AEs.

All the AEs (n=5) related to the investigational device were associated with skin and subcutaneous tissue disorders (primarily skin irritation). For four AEs that were related to the investigational device, the intensity was considered moderate, whereas one AE was considered severe (contact dermatitis).

Hence, it can be concluded that the current investigational device showed no unanticipated AEs considering that the subjects tried out a new type of product with different adhesive area and material.

7.7 Risk/ Benefit

The investigational device Heylo™ is a low-risk medical device (Class I) and there have been no unacceptable remaining risks identified related to its use through the risk management process.

AEs observed in the present investigation, that were possibly related, probably related or related to the investigational device such as skin irritations are known side effects of the use of an ostomy solution and the related risk of experiencing leakage when having an ostomy.

The number of AEs observed during this investigation were similar to the prevalence observed with the use of other ostomy solutions.

Contrary, several immediately observed benefits to patients' QoL and participation in everyday- and social activities were shown with its use. Therefore, the overall benefits with the use of the investigational device exceed the risks with its use to a very high degree.

7.8 Conclusion

This randomized controlled cross-over trial demonstrated that after 8 weeks of product usage, the test device Heylo™ provided positive care effects to QoL and the overall burden of living with a stoma, reflected as less embarrassment, better sleep, living with dignity and better capability of participating in society and interacting with close family and friends.

More specifically Heylo™ demonstrated significant improvements in all three OLI domain scores after 8 weeks product usage and in 5 out of 6 WHODAS 2.0 disability domain scores, together with an overall improvement in the total WHODAS 2.0 score and a reduced number of days with difficulties present. Also feeling of security and episodes of leakages outside the baseplate improved significantly after 8 weeks with Heylo™ compared to standard of care. Together, these findings suggest that Heylo™ provides clinically relevant and meaningful positive care effects for people living with a stoma. This intriguing finding was seen in study participants with both ileostomy and colostomy.

8 References

1. Martins L, Andersen BD, Colwell J, Down G, Forest-Lalande L, Novakova S, et al. Challenges faced by people with a stoma: peristomal body profile risk factors and leakage. British Journal of Nursing. 2022;31(7):376-85.
2. Claessens I, Probert R, Tielemans C, Steen A, Nilsson C, Andersen BD, et al. The Ostomy Life Study: the everyday challenges faced by people living with a stoma in a snapshot. Gastrointestinal Nursing. 2015;13(5):18-25.
3. Jeppesen PB, Vestergaard M, Boisen EB, Ajslev TA. Impact of Stoma Leakage in Everyday Life: Data from the Ostomy Life Study 2019. Br J Nurs. 2022;31:S48–S58.
4. Hedegaard C, Ajslev T, Zeeberg R, Steen Hansen A. Leakage and peristomal skin complications influences user comfort and confidence and are associated with reduced quality of life in people with a stoma. WCET® Journal. 2020;40(Epub ahead of print).
5. Davis JS, Svavarsdóttir MH, Pudło M, Arena R, Lee Y, Jensen MK. Factors impairing quality of life for people with an ostomy. Gastrointestinal Nursing. 2011;9(Sup2):14-8.
6. Rolls N, Yssing C, Bøgelund M, Håkan-Bloch J, de Fries Jensen L. Utilities associated with stoma-related complications: peristomal skin complications and leakages. Journal of Medical Economics. 2022;25(1):1005-14.
7. de Fries Jensen L, Rolls N, Russell-Roberts P, Vestergaard M, Jensen ML, Boisen EB. Leakage of stomal effluent outside the baseplate leads to rise in product usage and health professional interactions. British Journal of Nursing. 2023;32(1):8-19.
8. Coloplast AS. Clinical Investigation Report - CP278. Veeva. 2019;VV-0264051.
9. Coloplast AS. Clinical Investigation Report CP308. Veeva. 2021;VV-0313951.
10. Coloplast AS. Clinical Investigation Report CP321. Veeva. 2022;VV-0367954.
11. Coloplast AS. Clinical Investigation Report CP340. Veeva. 2022;VV-0386865.
12. Clinical Evaluation Report Heylo. VV-0315139. Coloplast A/S. 2021.
13. Coloplast AS. IFU master approved for CE mark. Veeva. 2021;VV-0331823.
14. Coloplast AS. IFU Master Heylo sensor layer for CE mark. Veeva. 2021;VV-0331824.
15. Coloplast AS. Clinical Investigation Plan CP345. Veeva. 2021;VV-0336576.
16. Nafees B, Storling ZM, Hindsberger C, Lloyd A. The ostomy leak impact tool: development and validation of a new patient-reported tool to measure the burden of leakage in ostomy device users. Health Qual Life Outcomes. 2018;16(1):231.
17. World Health Organization W. Measuring health and disability : manual for WHO Disability Assessment Schedule (WHODAS 2.0). Ustun TB, Kostanjesek N, Chatterji S, Rehm J, Organization WH, editors. World Health Organization: World Health Organization; 2010. 88 p.
18. Ustün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. Bull World Health Organ. 2010;88(11):815-23.
19. Coloplast AS. Ostomy Leak Impact Tool German (DE) - CP345. Veeva. 2023;VV-0389744.
20. Federici S, Bracalenti M, Meloni F, Luciano JV. World Health Organization disability assessment schedule 2.0: An international systematic review. Disability and rehabilitation. 2017;39(23):2347-80.
21. Bejer A, Ćwirlej-Sozańska A, Wiśniowska-Szurlej A, Wilmowska-Pietruszyńska A, Spalek R, de Sire A, et al. Psychometric properties of the Polish version of the 36-item WHODAS 2.0 in patients with hip and knee osteoarthritis. Qual Life Res. 2021;30(8):2415-27.
22. Ćwirlej-Sozańska A, Bejer A, Wiśniowska-Szurlej A, Wilmowska-Pietruszyńska A, de Sire A, Spalek R, et al. Psychometric Properties of the Polish Version of the 36-Item WHODAS 2.0 in Patients with Low Back Pain. Int J Environ Res Public Health. 2020;17(19).
23. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res. 2011;11(2):171-84.
24. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61(2):102-9.

25. Shulman MA, Kasza J, Myles PS. Defining the Minimal Clinically Important Difference and Patient-acceptable Symptom State Score for Disability Assessment in Surgical Patients. Anesthesiology. 2020;132(6):1362-70.
26. Wong JJ, DeSouza A, Hogg-Johnson S, De Groote W, Southerst D, Belchos M, et al. Measurement Properties and Minimal Important Change of the World Health Organization Disability Assessment Schedule 2.0 in Persons With Low Back Pain: A Systematic Review. Arch Phys Med Rehabil. 2022.
27. Pösl M, Cieza A, Stucki G. Psychometric properties of the WHODASII in rehabilitation patients. Quality of Life Research. 2007;16:1521-31.
28. Coloplast AS. Data Evaluability Form - CP345. Veeva. 2022;VV-0336595.
29. Coloplast AS. Statistical Analysis Plan - CP345. Veeva. 2022;VV-0386658.
30. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical psychology review. 1988;8(1):77-100.
31. Nafees B, Størling ZM, Hindsberger C, Lloyd A. The ostomy leak impact tool: development and validation of a new patient-reported tool to measure the burden of leakage in ostomy device users. Health Qual Life Outcomes. 2018;16(231).
32. Martins L, Andersen BD, Colwell J, Down G, Forest-Lalande L, Novakova S, et al. Challenges faced by people with a stoma: peristomal body profile risk factors and leakage. Br J Nurs. 2022;31(7):376-85.
33. Lee HH, Shin E-K, Shin H-I, Yang EJ. Is WHODAS 2.0 useful for colorectal cancer survivors? Ann Rehabil Med. 2017;41(4):667-76.
34. Coloplast AS. Statistical Analysis Report - Listings - CP345. Veeva. 2023;VV-0388333.
35. Coloplast AS. Notes to File_20221111_Questionnaires out of window_GE_CP345. Veeva. 2022;VV-0385907.
36. Coloplast AS. Notes to File_20221111_Informed Consent_GE_CP345. Veeva. 2022;VV-0385908.
37. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in psychology. 2013;4:863.
38. Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. Journal of graduate medical education. 2012;4(3):279-82.
39. Lee HH, Shin EK, Shin HI, Yang EJ. Is WHODAS 2.0 Useful for Colorectal Cancer Survivors? Ann Rehabil Med. 2017;41(4):667-76.
40. Coloplast AS. Statistical Analysis Report - Tables - CP345. Veeva. 2023;VV-0388334.
41. Jemec GB, Martins L, Claessens I, Ayello EA, Hansen AS, Poulsen LH, et al. Assessing peristomal skin changes in ostomy patients: validation of the Ostomy Skin Tool. Br J Dermatol. 2011;164(2):330-5.
42. Schott LL, Eaves D, Inglese G, Sinha M. Characteristics, Hospital Length of Stay, and Readmissions Among Individuals Undergoing Abdominal Ostomy Surgery: Review of a Large US Healthcare Database. Journal of Wound, Ostomy and Continence Nursing. 2022;49(6):529-39.
43. Colwell JC, Pittman J, Raizman R, Salvadalena G. A randomized controlled trial determining variances in ostomy skin conditions and the economic impact (ADVOCATE Trial). J Wound Ostomy Continence Nurs. 2018;45(1):37.
44. Lehur P-A, Deguines J-B, Montagliani L, Duffas J-P, Bresler L, Mauvais F, et al. Innovative appliance for colostomy patients: an interventional prospective pilot study. Tech Coloproctol. 2019;23(9):853-9.
45. Rat P, Robert N, Fernandes I, Edmond D, Mauvais F. Evaluation of a one-piece soft convex ostomy appliance: a prospective, multicentre, open-label pilot study. Br J Nurs. 2018;27(16):S20-S8.
46. Coloplast AS. Notes to File_20220124_Product Accountability_GE_CP345 Veeva. 2023;VV-0389202
47. Coloplast AS. Statistical Analysis Report - Figures - CP345. Veeva. 2023;VV-0388332.

9 Investigators and administrative structure of clinical investigation

9.1 List of investigators

PRINCIPAL INVESTIGATOR	CONTACT INFORMATION
[redacted]	[redacted]

Updated CVs were obtained for the principal investigators and study nurses (available in Coloplast A/S internal system).

9.2 Sponsor representatives

COORDINATING CLINICAL MANAGER	PRINCIPAL BIOSTATISTICIAN
[redacted] Clinical Operations, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]	[redacted] Clinical strategies, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]
PRINCIPAL MEDICAL AFFAIRS PROJECT MANAGER	**DATA MANAGER**
[redacted] Scientific Affairs, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]	[redacted] Clinical Operations, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]
DIRECTOR OF CLINICAL STRATEGIES	**DIRECTOR OF CLINICAL OPERATIONS**
[redacted] Clinical Strategies, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]	[redacted] Clinical Operations, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]

SENIOR MEDICAL WRITER	CLINICAL STRATEGY PROJECT MANAGER/ FORMER MEDICAL WRITER (AUTHOR/APPROVER OF VERSION 1.0)
[redacted] Clinical Strategies, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]	[redacted] Clinical Strategies, Payers & Evidence Coloplast A/S Holtedam 1-3 3050 Humlebæk Denmark [redacted]
DIRECTOR MEDICAL AFFAIRS, GERMANY	**CLINICAL LEAD, GERMANY**
[redacted] Medical Affairs Coloplast GmbH Am Neumarkt 42 22041 Hamburg Germany [redacted]	[redacted] Regulatory Affairs/Clinical Trials Coloplast GmbH Am Neumarkt 42 22041 Hamburg Germany [redacted]

9.3 Other

Not applicable.

10 Signature page

By signing this CIR electronically, the relevant parties (Sponsor) declare to have either prepared or read this report and confirm to the best of their knowledge that it accurately describes the conduct and results of the trial.

Investigators' agreement with the contents of this CIR are signed in an affidavit filed at Coloplast A/S: Affidavit Clinical Investigation Report - CP345, VV-0336571.

11 Change Log

Version no.	Initials	Short description of and reason for change
1.0	DKJOAT Feb 2023	Document established based on Clinical Investigation Report template, version 5.0.
2.0	DKMVES Jan 2024	Document has been updated with additional summary statistics and analyses as requested by BfArM. These additions include: Updated Figure 6 with information on timing of visits. Updates to text in section 6.3 and in Table 6, with height, weight and work status information. Update to statistical method section 5.3.20.7 with responder analyses for the primary and secondary endpoint and with subgroup analysis by age and gender for the secondary endpoint.

		Update to post-hoc sensitivity analysis section 6.9.2.5 with subgroup analysis for the secondary endpoint and with responder analysis for the primary and secondary endpoint. Added appendix 12.17, Summary statistics of drop-outs, which include two tables.
3.0	DKTEAA Feb. 2024	Updated affiliation of Prof. Dr Med P. Ambe and added few editorial changes.
4.0	DKMVES March 2024	This document has been updated with additional summary statistics and analyses as requested by BfArM. These additions include: Clarification several places that the primary efficacy estimates were tested after 8 weeks treatment on either Heylo™ or standard of care. Further, that a statistical significance level of $p<0.025$ for either the primary or secondary endpoint was considered sufficient on their own (as both endpoints were tested on equal terms). Added Post-hoc analyses for the primary and secondary endpoint considering all randomized patients with use of two different imputation models and including baseline values. A detailed description of the statistics has been included in Appendix 12.17. Table 6 of baseline summaries has been extended to include two additional sub-groups (treatment sequence A and treatment sequence B). Table 1 and 2 in Appendix 12.17 (related to drop-outs), were moved into the main text and SDs added. A table of efficacy analyses including standardized effect measures (Cohen's *d*'s) has been added. Results of responder analyses have been added in tabular form. Responder summaries for the primary endpoint and secondary endpoint have been added in tabular form and related Sankey diagrams have been added to the post-hoc result section, and discussion. Two sponsor representatives (DKMVES and DKJOAT) have been added to the list of study responsible.

12 Appendixes to the report

12.1 CIP and amendments

Clinical Investigation Plan CP345; Veeva no: VV-0336576 (15).

12.2 The Ethic Committee

Approval site 1
Ethik-Kommission der Universität Witten/Herdecke e.V. z. Hd. Herrn RA Prof. Dr. med. P. Gaidzik Alfred-Herrhausen-Straße 50 58448 Witten Sekretariat: Frau Andrea Pleger Mo-Fr 8.00-12.00 Uhr Telefon 02302/926-740 Telefax 02302/926-739 e-mail: sekretariat-ethik@uni-wh.de Internet: www.ethik-kommission-uwh.de

12.3 Evaluability of subjects

The evaluability form is available in Coloplast internal document handling system: Data Evaluability Form - CP345, Veeva no: VV-0336595 (28).

12.4 Tables, figures and graphs referred to but not included in the text

Not applicable.

12.5 Product accountability

Product accountability logs have been maintained. Some issues and have been identified. These issues are documented in the note-to-file VV-0389202 (46).

12.6 Statistical plan and report

The statistical plan and report are available in Coloplast internal document handling system.

- Statistical Analysis Plan CP345, VV-0386658 (29)
- Statistical Analysis Report - Figures - CP345, VV-0388332 (47)
- Statistical Analysis Report - Tables - CP345, VV-0388334 (40)
- Statistical Analysis Report - Listings - CP345 VV-0388333 (34)

12.7 Definitions of adverse events, adverse device effects and device deficiencies

Device deficiency
A device deficiency refers to the inadequacy of the investigational medical device with respect to its identity, quality, durability, reliability, safety or performance. This includes malfunctions, misuse or use errors and inadequate labelling.

Primary and secondary endpoints measured during this investigation were not reported as device deficiencies.

Adverse event
An AE is any untoward medical occurrence, unintended disease or injury, or untoward clinical signs (including abnormal laboratory findings) that may occur in subjects, users or other parties, whether or not it is related to the medical device(s), or the procedures involved. This could include events such as headache or dizziness.

Adverse device effect
An AE, which is related to the use of the investigational medical device, is an ADE, and should be marked as related or possibly related to the medical devices(s) on the AE form.

The definition of an ADE includes any event resulting from insufficiencies or inadequacies in the instruction for use, malfunction of the device, use error or from intentional misuse of the device, deployment, implantation, installation and operation.

Serious adverse event
A serious adverse event is an adverse event that:

- Led to death,
- Led to a serious deterioration in health of the subject that either resulted in:
    1) a life-threatening illness or injury, or
    2) a permanent impairment of a body structure or a body function, or
    3) required inpatient hospitalization or prolongation of existing hospitalization, or
    4) medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function.
- Led to fetal distress, fetal death or a congenital abnormality or birth defect.

This includes device deficiencies that might have led to a serious adverse event if:

- Suitable action had not been taken, or
- Intervention had not been made, or
- Circumstances had been less fortunate.

These were handled under the serious adverse event reporting.

Planned hospitalization for a pre-existing condition, or a procedure required by the CIP, without serious deterioration in health, is not considered an SAE.

Serious adverse device effect
A SADE is an ADE that results in any of the consequences characteristic of a SAE.
A SADE may be an ASADE or a USADE.

Anticipated serious adverse device effect
An ASADE is any event that by its nature, incidence, severity or outcome has been previously identified in the risk analysis report.

Unanticipated serious adverse device effect
An USADE is a SADE which by its nature, incidence, severity or outcome has not been identified in the current version of the risk analysis report.

12.8Endpoints and assessment flowchart

	Endpoints and assessment	Assessed by	V1 Baseline	V2	V3	Every 2nd week	Every 4th week
Primary endpoint	*Emotional impact* score OLI scale – Appendix 12.9	Subject	X	X	X	X	
Secondary endpoint	Participation in society domain score WHODAS 2.0 – Appendix 12.10	Subject	X	X	X		X
Exploratory endpoints	Impact on coping and in control score OLI scale – Appendix 12.9	Subject	X	X	X	X	
	Impact on usual and social activities score OLI scale – Appendix 12.9	Subject	X	X	X	X	
	Feeling of security - Question: "*How was the feeling of security while wearing the product?"* Answers: Very poor/Poor/Acceptable/Good/Very good.	Subject	X	X	X		
	Cognition domain score Mobility domain score Self-care domain score Getting along with people domain score Life activities (household and work) domain score WHODAS 2.0 – Appendix 12.10	Subject	X	X	X		X
Assessments	Leakage outside baseplate – Question: *"Think back on the last 2 weeks; how many times have you experienced stoma effluent leakage outside the baseplate (e.g. onto clothes or bedsheets)?" (number)*	Subject	X	X	X	X	
	Change in current stoma product – *Question: "Has there been any change in current stoma product during the test period? (Yes/No) If yes, please add: Type (1pc/2pc), Kind (Flat, Convex, Concave), Brand (Coloplast, Hollister, Dansac, Salts, other)"*	Investigator		X	X		
	Change in Heylo size – Question: "*Change of Heylo size needed?" (Yes/no), if yes, please provide the new size: 40 mm, 50 mm, 60 mm, 70 mm, 80 mm"*	Investigator		X	X		
	Adverse events/device deficiencies	Investigator	X	X	X		

12.9 Ostomy Leakage Impact tool and Linguistic validation certificate

Emotional impact

When you thought about your ostomy device and the risk of leakage, what emotions did you feel?

In the last 7 days, due to leakage or worry about leakage…	**All of the time**	**Often**	**Sometimes**	**Rarely or never**
1. I felt panic	0	1	2	3
2. I felt stressed out	0	1	2	3
3. I felt more afraid about leaks in the future	0	1	2	3
4. I felt worry	0	1	2	3
5. I felt frustrated	0	1	2	3
6. I felt embarrassed	0	1	2	3
7. I felt worried that I might leak	0	1	2	3
8. I couldn't sleep	0	1	2	3
9. I kept waking up at night to check my stoma	0	1	2	3
10. I kept checking my ostomy bag to see if I have leaked	0	1	2	3

Usual and Social activities

When you thought about your ostomy device and the risk of leakage, how did it affect your activities?

In the last 7 days due to leakage or worry about leakage…	**All of the time**	**Often**	**Sometimes**	**Rarely or never**	***Not applicable***
11. I decided to stay at home	0	1	2	3	9
12. I couldn't do light activities	0	1	2	3	9
13. I changed my plans	0	1	2	3	9
14. I was unable to go out and meet family and friends	0	1	2	3	9
15. I avoided close physical contact with family and friends	0	1	2	3	9
16. I did not want to see people	0	1	2	3	9
17. I avoided people	0	1	2	3	9
18. I tried to avoid meeting new people	0	1	2	3	9

Coping and in control

When you thought about your ostomy device and the risk of leakage, how did it affect your ability to cope?

In the last 7 days, due to leakage or worry about leakage…	**All of the time**	**Often**	**Sometimes**	**Rarely or never**
19. I felt in control	0	1	2	3
20. I was able to cope	0	1	2	3
21. I felt calm	0	1	2	3
22. I saw my friends as I usually do	0	1	2	3

ICON LANGUAGE SERVICES
A Symbol of Excellence

LINGUISTIC VALIDATION CERTIFICATE

Ostomy Leak Impact Tool

This is to certify that **ICON Language Services** conducted the linguistic validation of the **paper** version of the **Ostomy Leak Impact Tool** (source file name: **Ostomy Leak Impact Tool - German**) into **German for Germany**.

The aim of linguistic validation is to obtain translations that are:

- conceptually equivalent to the original and comparable across languages;
- culturally relevant to the context of the target country;
- easily understood by the people to whom the translated instrument is administered.

This is achieved using a rigorous ISO-17100 certified methodology[1] involving:

- a process which comprises several steps;
- the instrument's developer input on conceptual issues;
- a skilled team recruited by **ICON Language Services** in the target country and headed by a consultant with knowledge of and experience in the field of Clinical Outcome Assessments. The consultant supervises and coordinates the linguistic validation process in his/her country;
- a centralized review process coordinated by **ICON Language Services**, including quality control by linguists and discussions about translation decisions with the consultant at each step of the process.
- cross-cultural harmonisation to ensure common understanding of the instrument's concepts by all participants involved in the process and achieve conceptual equivalence across languages.

The aforementioned translation (filename : Ostomy Leak Impact Tool_deu-DE_19AUG2021, dated 19 August 2021) underwent the following steps:

- Quality Check through a backtranslation
- Cognitive Interview step on 5 patients who live with a stoma
- Proofreading step

The exiting version on which the linguistic validation work was based has not been produced by ICON.

ICON Language Services may not be held liable for any changes made on the translation after completion of project **0649-TR-0005** by **ICON Language Services** on **19 August 2021**.

Charlotte Traynor
19 Aug 2021 14:40:019+0000

REASON: I approve this document

306df1b2-f2e2-4824-8b97-b6597c94b09a

Charlotte Traynor
Project Manager
ICON Language Services

[1] **References:**
- Acquadro C., Jambon B., Ellis D. and Marquis P. Language and translation issues. In Spilker B, ed. Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia: Lippincott-Raven Publishers, 1996: 575-585.
- Linguistic Validation Manual for Health Outcomes Assessments. Acquadro C, Conway K, Giroudet C, Mear I. Second Edition - Mapi Institute, Lyon, France, January 2012 - ISBN: 2-9522021-0-9

Ostomy Leak Impact Tool – Germany/German - Version of 19 Aug 2021 – ICON.
ID0649-TR-0005 / Ostomy Leak Impact Tool_deu-DE_19AUG2021.xlsx

Page | 1

DocUUID: 24e1eeb7-c655-49f1-9d15-f4dfd43892f4

12.10 WHODAS 2.0

WHODAS 2.0

WORLD HEALTH ORGANIZATION
DISABILITY ASSESSMENT SCHEDULE 2.0

36-item version, self-administered

This questionnaire asks about difficulties due to health conditions. Health conditions include diseases or illnesses, other health problems that may be short or long lasting, injuries, mental or emotional problems, and problems with alcohol or drugs.

Think back over the past 30 days and answer these questions, thinking about how much difficulty you had doing the following activities. For each question, please circle only one response.

In the past 30 days, how much difficulty did you have in:						
Understanding and communicating						
D1.1	Concentrating on doing something for ten minutes?	None	Mild	Moderate	Severe	Extreme or cannot do
D1.2	Remembering to do important things?	None	Mild	Moderate	Severe	Extreme or cannot do
D1.3	Analysing and finding solutions to problems in day-to-day life?	None	Mild	Moderate	Severe	Extreme or cannot do
D1.4	Learning a new task, for example, learning how to get to a new place?	None	Mild	Moderate	Severe	Extreme or cannot do
D1.5	Generally understanding what people say?	None	Mild	Moderate	Severe	Extreme or cannot do
D1.6	Starting and maintaining a conversation?	None	Mild	Moderate	Severe	Extreme or cannot do
Getting around						
D2.1	Standing for long periods such as 30 minutes?	None	Mild	Moderate	Severe	Extreme or cannot do
D2.2	Standing up from sitting down?	None	Mild	Moderate	Severe	Extreme or cannot do
D2.3	Moving around inside your home?	None	Mild	Moderate	Severe	Extreme or cannot do
D2.4	Getting out of your home?	None	Mild	Moderate	Severe	Extreme or cannot do
D2.5	Walking a long distance such as a kilometre [or equivalent]?	None	Mild	Moderate	Severe	Extreme or cannot do

Please continue to next page ...

WHODAS 2.0

WORLD HEALTH ORGANIZATION
DISABILITY ASSESSMENT SCHEDULE 2.0

36
Self

In the past 30 days, how much difficulty did you have in:						
Self-care						
D3.1	Washing your whole body?	None	Mild	Moderate	Severe	Extreme or cannot do
D3.2	Getting dressed?	None	Mild	Moderate	Severe	Extreme or cannot do
D3.3	Eating?	None	Mild	Moderate	Severe	Extreme or cannot do
D3.4	Staying by yourself for a few days?	None	Mild	Moderate	Severe	Extreme or cannot do
Getting along with people						
D4.1	Dealing with people you do not know?	None	Mild	Moderate	Severe	Extreme or cannot do
D4.2	Maintaining a friendship?	None	Mild	Moderate	Severe	Extreme or cannot do
D4.3	Getting along with people who are close to you?	None	Mild	Moderate	Severe	Extreme or cannot do
D4.4	Making new friends?	None	Mild	Moderate	Severe	Extreme or cannot do
D4.5	Sexual activities?	None	Mild	Moderate	Severe	Extreme or cannot do
Life activities						
D5.1	Taking care of your household responsibilities?	None	Mild	Moderate	Severe	Extreme or cannot do
D5.2	Doing most important household tasks well?	None	Mild	Moderate	Severe	Extreme or cannot do
D5.3	Getting all the household work done that you needed to do?	None	Mild	Moderate	Severe	Extreme or cannot do
D5.4	Getting your household work done as quickly as needed?	None	Mild	Moderate	Severe	Extreme or cannot do

Please continue to next page ...

If you work (paid, non-paid, self-employed) or go to school, complete questions D5.5–D5.8, below. Otherwise, skip to D6.1.

Because of your health condition, in the past 30 days, how much difficulty did you have in:						
D5.5	Your day-to-day work/school?	None	Mild	Moderate	Severe	Extreme or cannot do
D5.6	Doing your most important work/school tasks well?	None	Mild	Moderate	Severe	Extreme or cannot do
D5.7	Getting all the work done that you need to do?	None	Mild	Moderate	Severe	Extreme or cannot do
D5.8	Getting your work done as quickly as needed?	None	Mild	Moderate	Severe	Extreme or cannot do

Participation in society						
In the past 30 days:						
D6.1	How much of a problem did you have in joining in community activities (for example, festivities, religious or other activities) in the same way as anyone else can?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.2	How much of a problem did you have because of barriers or hindrances in the world around you?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.3	How much of a problem did you have living with dignity because of the attitudes and actions of others?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.4	How much time did you spend on your health condition, or its consequences?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.5	How much have you been emotionally affected by your health condition?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.6	How much has your health been a drain on the financial resources of you or your family?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.7	How much of a problem did your family have because of your health problems?	None	Mild	Moderate	Severe	Extreme or cannot do
D6.8	How much of a problem did you have in doing things by yourself for relaxation or pleasure?	None	Mild	Moderate	Severe	Extreme or cannot do

Please continue to next page ...

WHODAS 2.0
WORLD HEALTH ORGANIZATION
DISABILITY ASSESSMENT SCHEDULE 2.0

36
Self

H1	Overall, in the past 30 days, how many days were these difficulties present?	***Record number of days*** ____
H2	In the past 30 days, for how many days were you totally unable to carry out your usual activities or work because of any health condition?	***Record number of days*** ____
H3	In the past 30 days, not counting the days that you were totally unable, for how many days did you cut back or reduce your usual activities or work because of any health condition?	***Record number of days*** ____

This completes the questionnaire. Thank you.

12.11 Bristol Scale

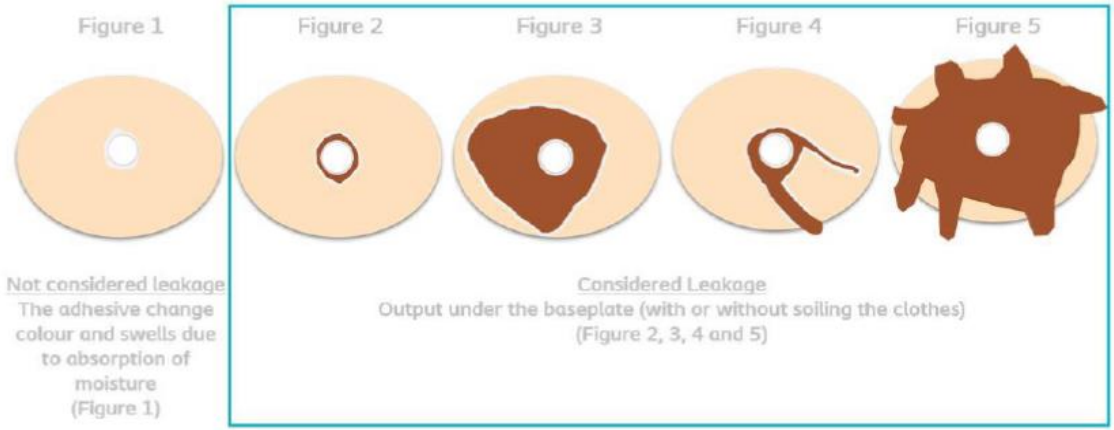
The Bristol stool form scale

The stool type illustrations below will help you determine your stool type.

Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
						
Separate hard lumps, like nuts (hard to pass)	Sausage-shaped but lumpy	Like a sausage but with cracks on its surface	Like a sausage or snake, smooth and soft	Soft blobs with clear-cut edges (passed easily)	Fluffy pieces with ragged edges, a mushy stool	Watery, no solid pieces, entirely liquid

Lewis SJ, Heaton KW (1997). "Stool form scale as a useful guide to intestinal transit time". Scand. J. Gastroenterol. 32 (9): 920–4

12.12 Definition of inclusion criteria: “Leakage defined as output/seeping under the baseplate”



12.13 Summary statistics over endpoints after 8 weeks

	Heylo™	Standard of care
OLI domains:		
Domain 1: Emotional impact (0-100)		
N	122	128
Mean (SD)	74.2 (25.7	61.4 (27.3)
Min;Max	0; 100	0; 100
Median	83.3	60.0
Domain 2: Usual and social activities (Scale 0-100)		
N	72*	89^
Mean (SD)	77.3 (25.6)	71.2 (24.3)
Min;Max	0; 100	0; 100
Median	87.5	70.8
Domain 3: Coping and in control (Scale 0-100)		
N	122	128
Mean (SD)	75.1 (30.0)	65.4 (29.9)
Min;Max	0; 100	0; 100
Median	91.7	66.7
WHODAS domain 1-6 and total score		
Domain 1: Cognition		
N	121	128
Mean (SD)	21.4 (21.3)	25.2 (22.3)
Min;Max	0; 80	0; 100
Median	15.0	25.0
Domain 2: Mobility		
N	121	128
Mean (SD)	26.1 (25.1)	30.5 (26.5)
Min;Max	0; 94	0; 94
Median	18.8	25.0
Domain 3: Self-care		
N	121	128
Mean (SD)	16.9 (22.3)	19.9 (22.6)
Min;Max	0; 90	0; 90
Median	10.0	10.0
Domain 4: Getting along		
N	121	128
Mean (SD)	27.2 (25.7)	31.0 (26.5)
Min;Max	0; 100	0; 100
Median	25.0	25.0
Domain 5 Life activities		
N	121	128
Mean (SD)	34.2 (28.0)	39.6 (29.1)
Min;Max	0; 100	0; 100
Median	33.3	40.0
Domain 6 Participation		
N	121	128
Mean (SD)	32.5 (24.8)	37.8 (23.8)
Min;Max	0; 92	0; 92
Median	29.2	37.5
Total WHODAS score		

N	121	128
Mean (SD)	26.4 (20.9)	30.7 (22.1)
Min;Max	0; 81	0; 85
Median	22.3	28.0
No of times with output onto clothes last 2 weeks?		
N	122	128
Mean (SD)	1.5 (2.1)	2.3 (2.7)
Min;Max	0; 14	0; 15
Median	1.0	2.0
How was the feeling of security while wearing the test product?		
N	121 (100.0)	128 (100.0)
Very poor	1 (0.8)	1 (0.8)
Poor	4 (3.3)	12 (9.4)
Acceptable	24 (19.8)	45 (35.2)
Good	50 (41.3)	47 (36.7)
Very good	42 (34.7)	23 (18.0)

*50 subjects responded ‘not applicable’ to the questions in this domain
^39 subjects responded ‘not applicable’ to the questions in this domain

12.14 Box plots for primary and secondary endpoints

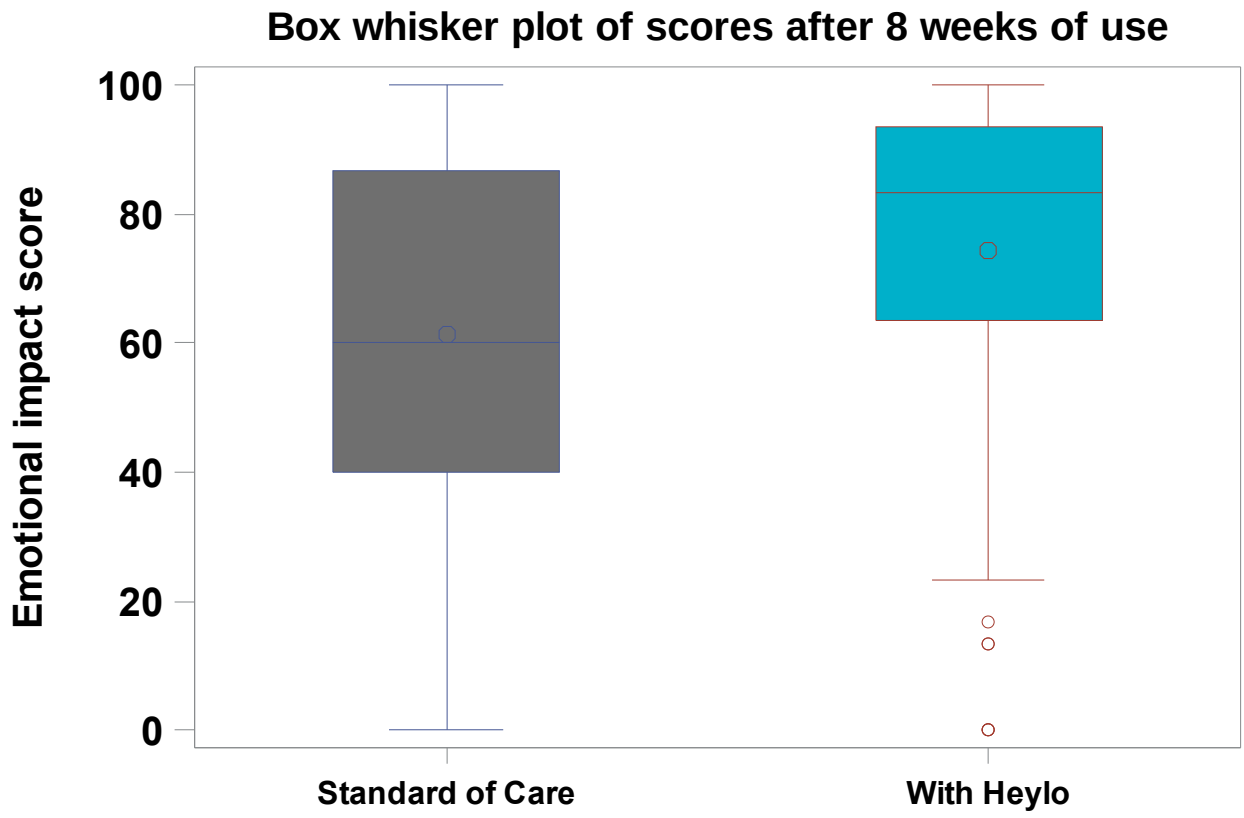


Figure 1 Box plot of the raw data for the primary endpoint (The Emotional impact score). The box within the box plot represents the interquartile range (IQR) which shows the middle 50% of the scores and is calculated by subtracting the lower 25% quartile from the upper 75% quartile. The line and the small circle within the box represent the median and the mean of the observations, respectively. The variation in data is presented by the upper and lower whisker. The upper whisker ends at the maximum observation below the upper fence, defined as 1.5*IQR above 75% quartile, whereas the lower whisker ends at the minimum observation above the lower fence defined as 1.5*IQR below 25% quartile. Observations beyond the whiskers are plotted as individual small circles.

Figure 2 Box plot of the raw data for the secondary endpoint (The *Participation* domain score). The box within the box plot represents the interquartile range (IQR) which shows the middle 50% of the scores and is calculated by subtracting the lower 25% quartile from the upper 75% quartile. The line and the small circle within the box represent the median and the mean of the observations, respectively. The variation in data is presented by the upper and lower whisker. The upper whisker ends at the maximum observation below the upper fence, defined as 1.5*IQR above 75% quartile, whereas the lower whisker ends at the minimum observation above the lower fence defined as 1.5*IQR below 25% quartile. Observations beyond the whiskers are plotted as individual small circles.

12.15 Plots to check for normal distribution of the residuals from the linear mixed models

Conditional residuals from the linear mixed model for the primary endpoint (the *Emotional impact* score (Leakage_emotion)) and for the secondary endpoint (the *Participation* domain 6 score (Whodas_do6)).

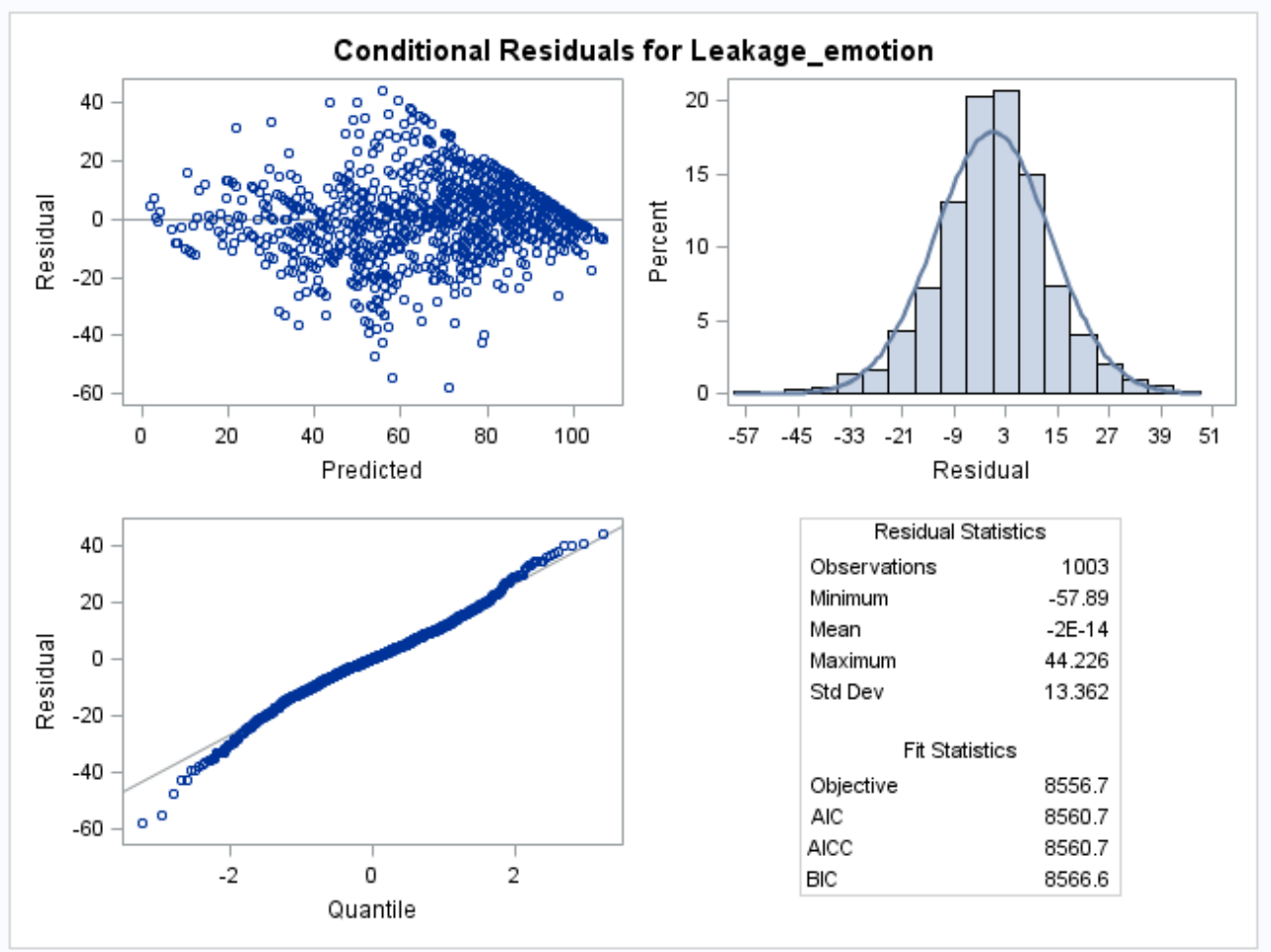


Figure 1 The panel consist of a scatterplot of residuals, a histogram with normal density, a Q-Q plot and summary statistics for the residuals and the model fit for the primary endpoint (the *Emotional* impact score).

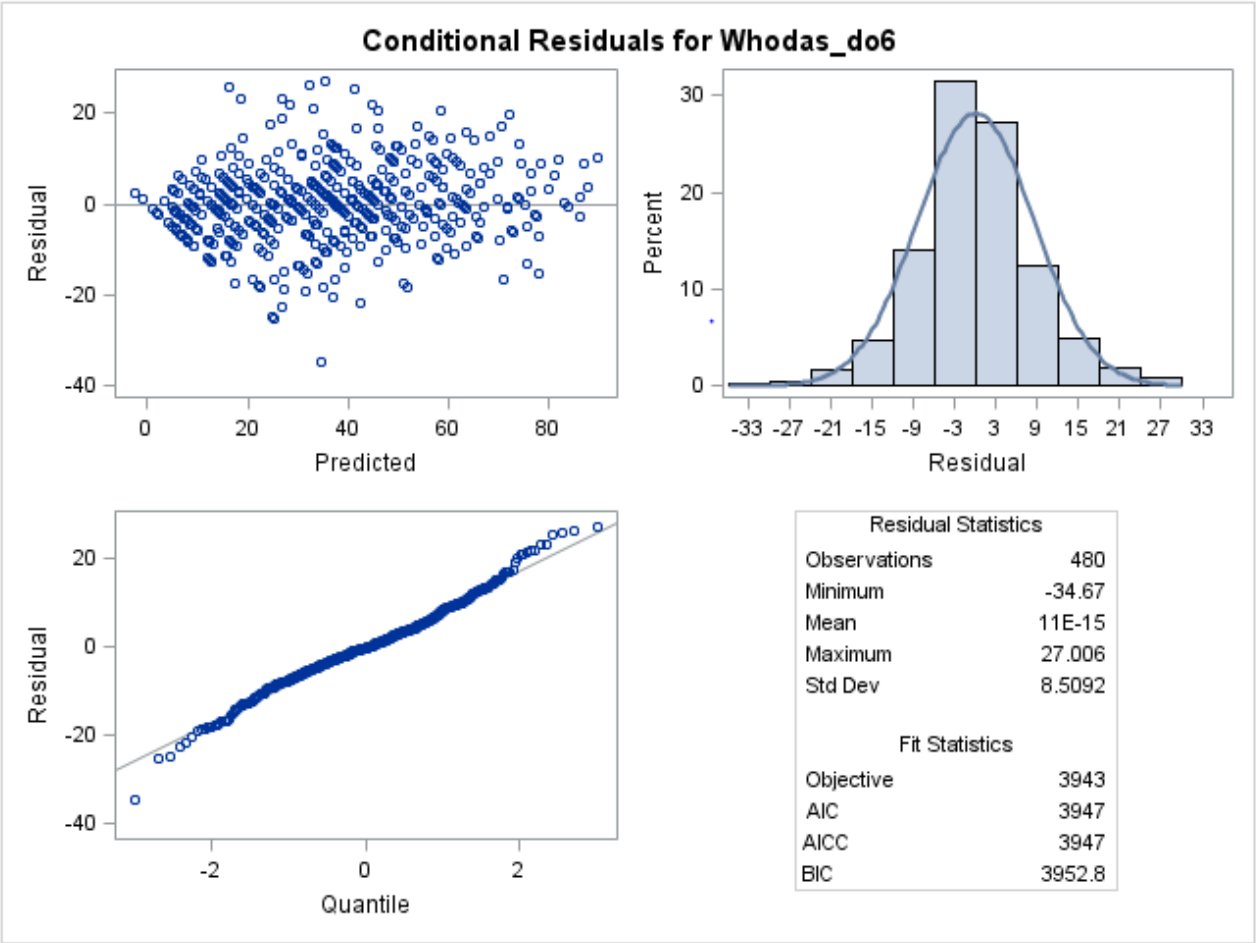


Figure 2 The panel consist of a scatterplot of residuals, a histogram with normal density, a Q-Q plot and summary statistics for the residuals and the model fit for the secondary endpoint (the *Participation* domain score).

12.16 Time trend evaluations

Time trend figures for the three OLI domain scores (Figure 1-3):

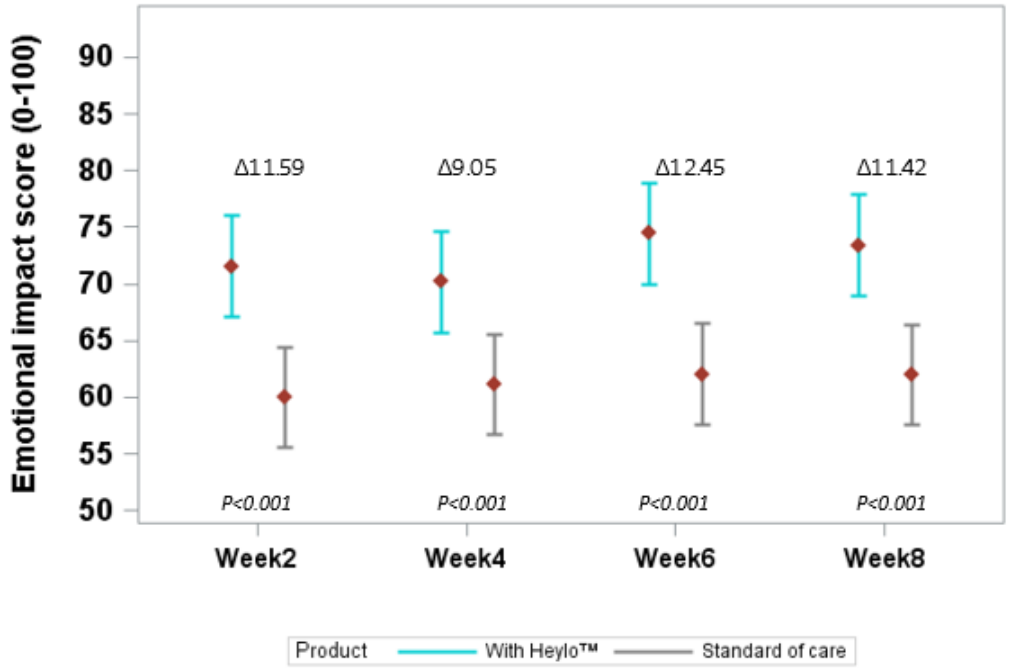


Figure 1 Ostomy leak impact (OLI) scores for the *Emotional impact* domain during the investigation. LS means including 95% confidence intervals and P-values for the Ostomy leak impact (OLI) score for Emotional impact domain scores during the investigation for Heylo™ and standard of care.

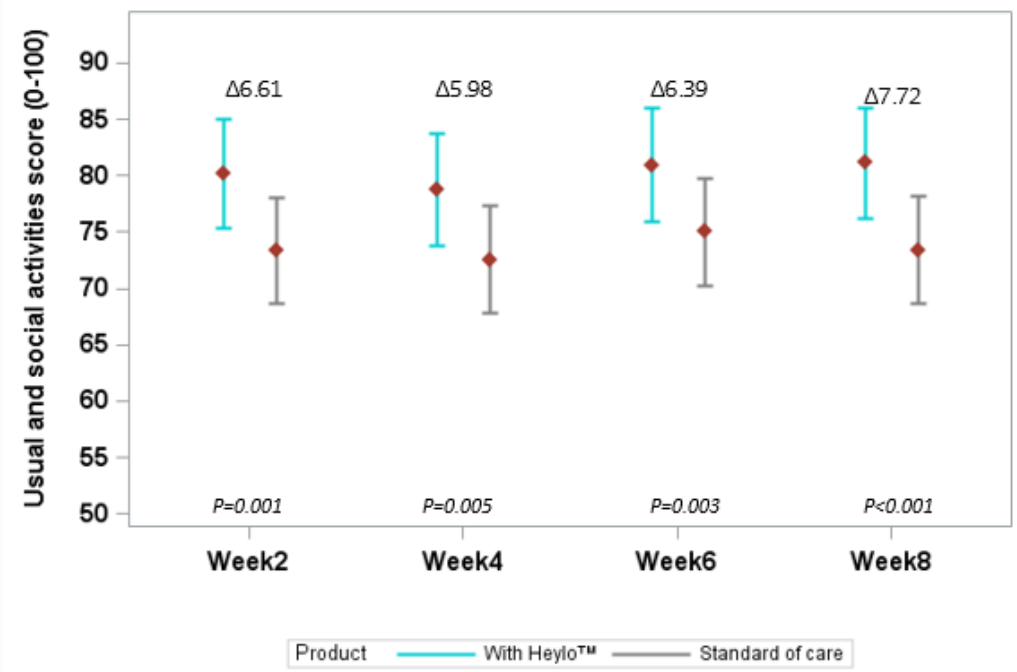


Figure 2 Ostomy leak impact (OLI) scores for the domain Usual and social activities during the investigation. LS means including 95% confidence intervals and P-values for the Ostomy leak impact (OLI) domain score for Usual and social activities for Heylo™ and standard of care during the investigation are displayed.

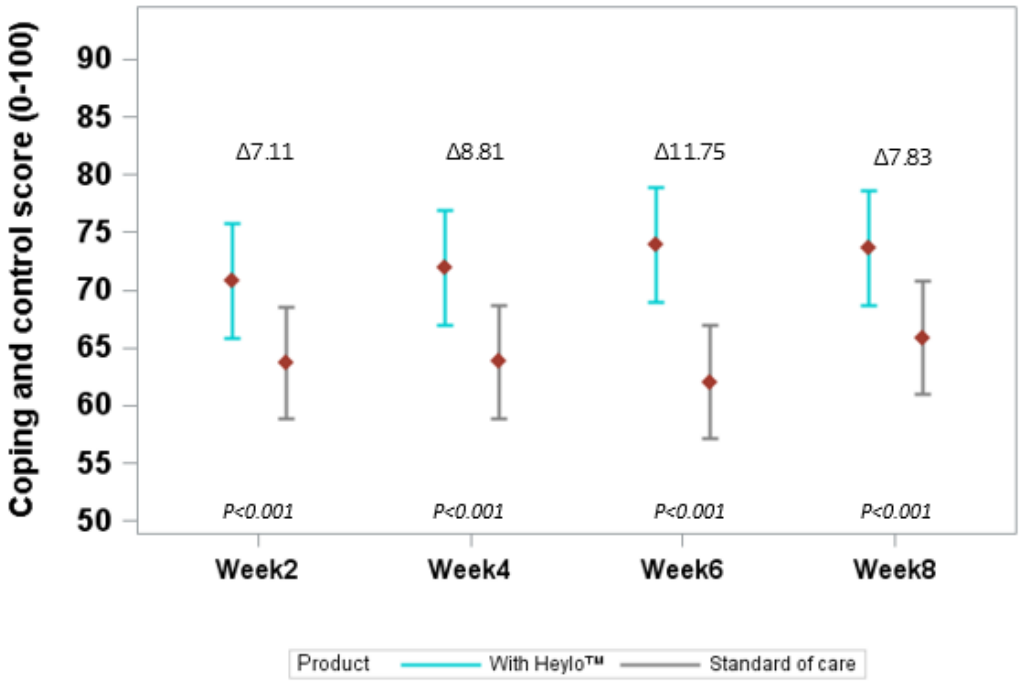


Figure 3 Ostomy leak impact (OLI) scores for the domain *Coping and in control*. LS means including 95% confidence intervals and P-values for the Ostomy leak impact (OLI) domain score for Coping and in control for Heylo™ and standard of care during the investigation are displayed.

WHODAS 2.0 domain (D1-D6) scores and total WHODAS 2.0 score during the investigation (Figure 4-10):

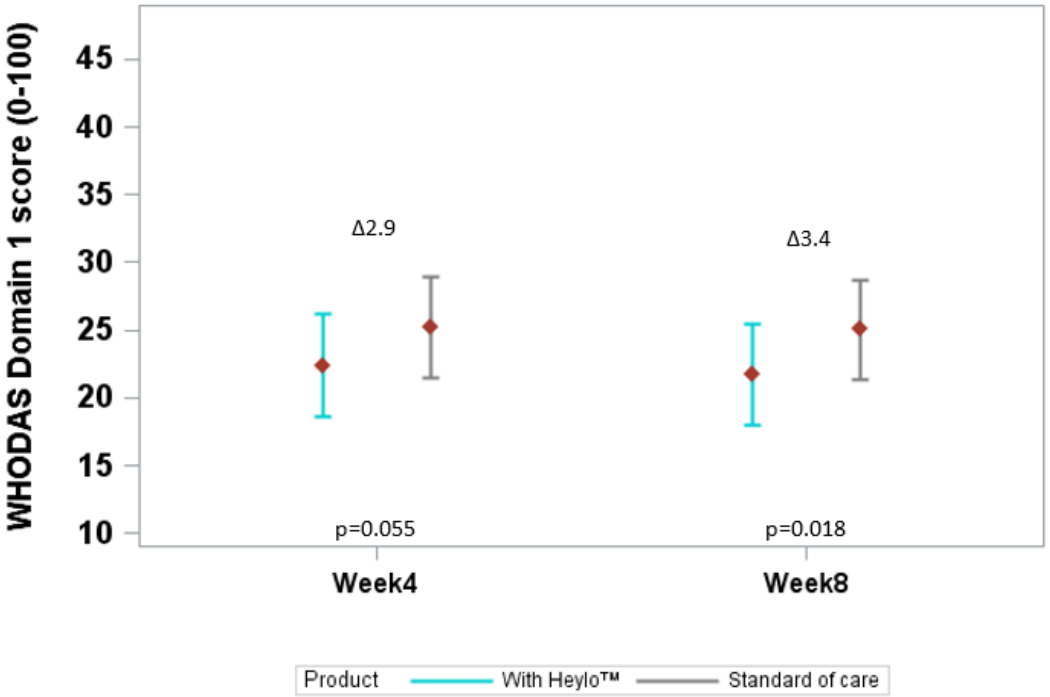


Figure 4 WHODAS 2.0 domain 1 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

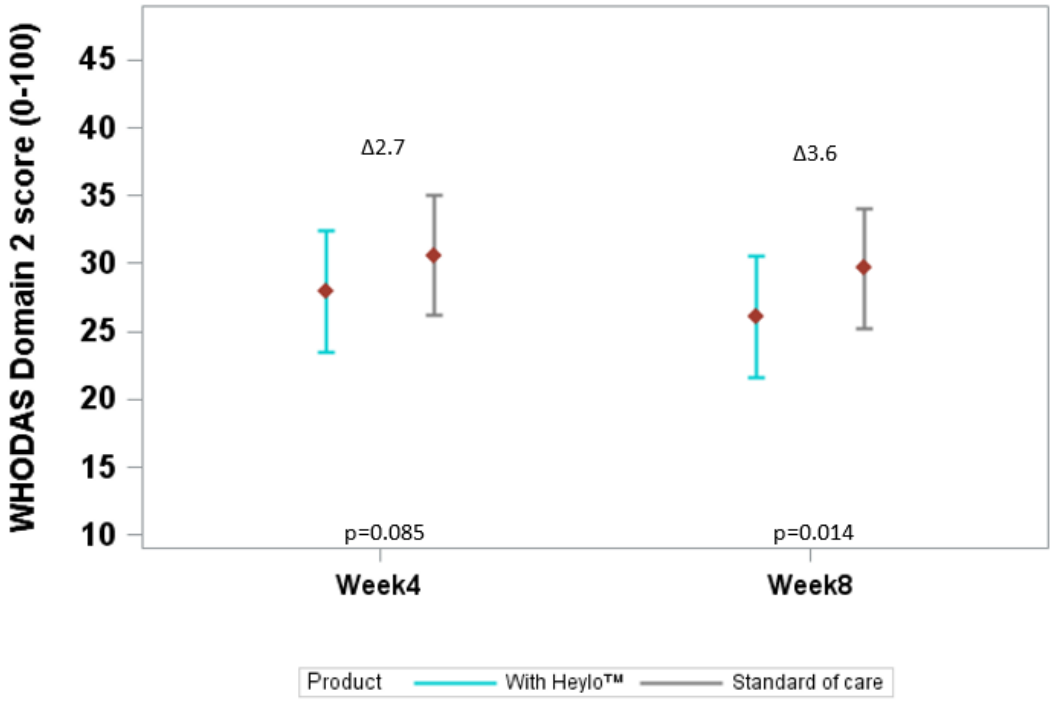


Figure 5 WHODAS 2.0 domain 2 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

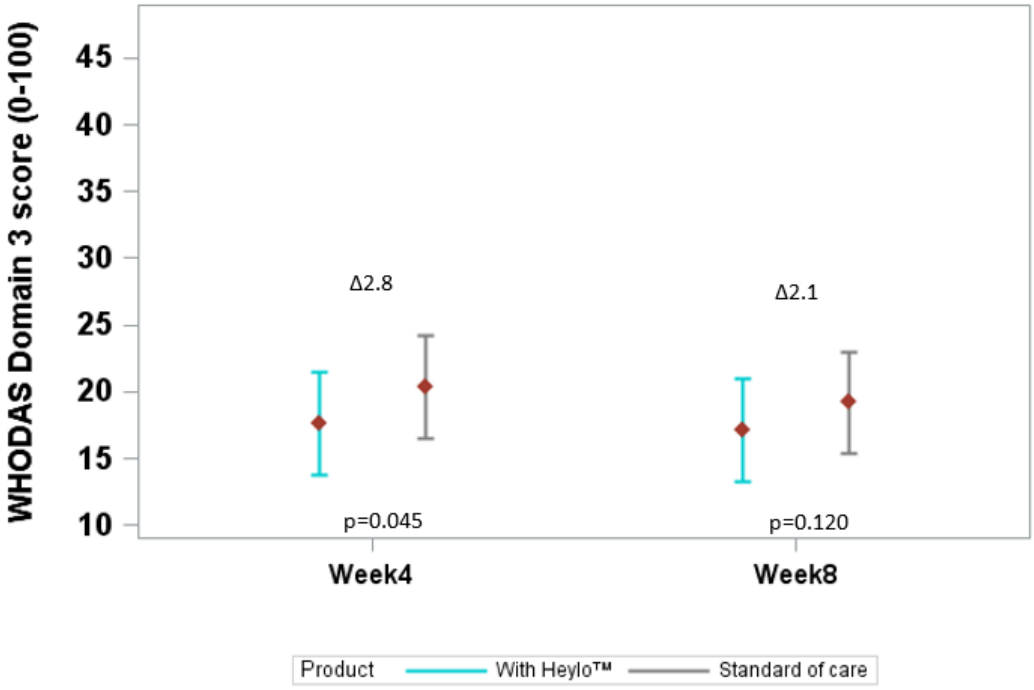


Figure 6 WHODAS 2.0 domain 3 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

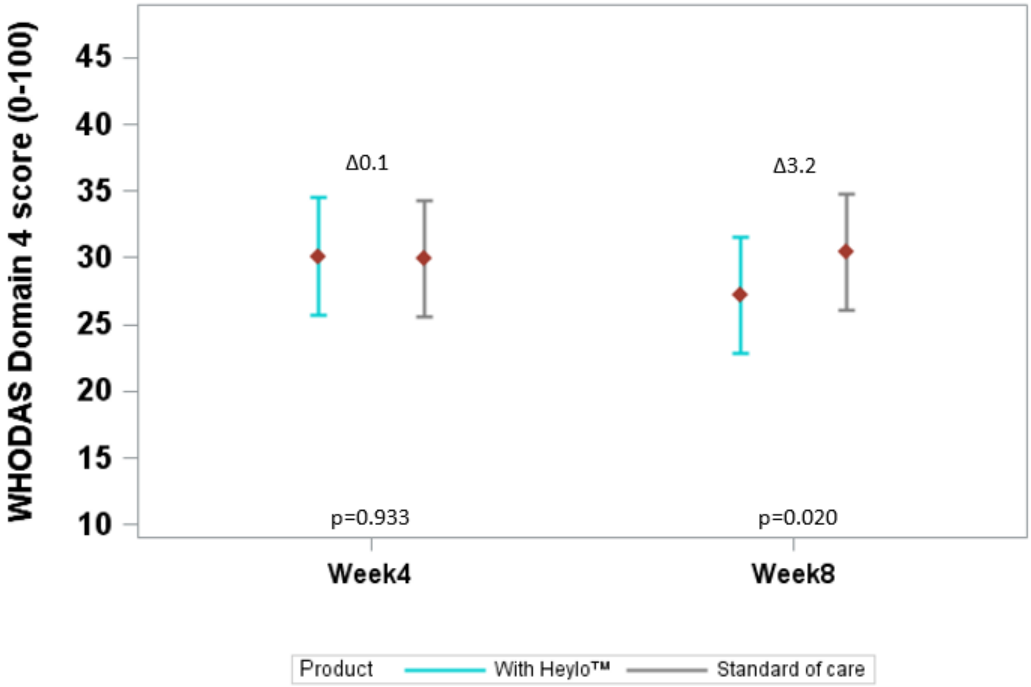


Figure 7 WHODAS 2.0 domain 4 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

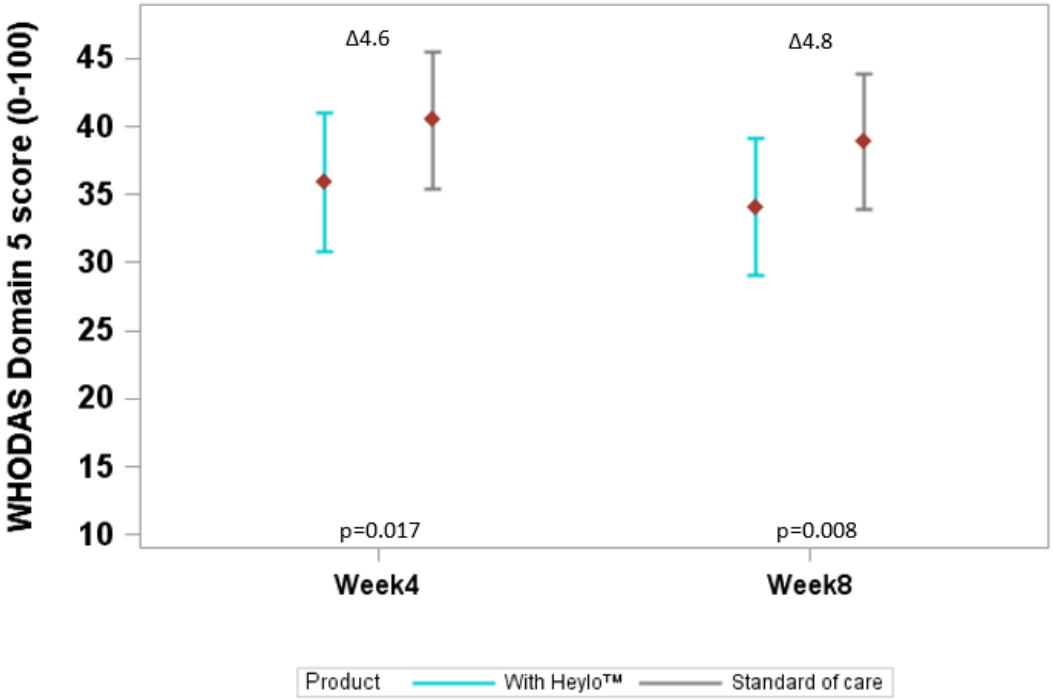


Figure 8 WHODAS 2.0 domain 5 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

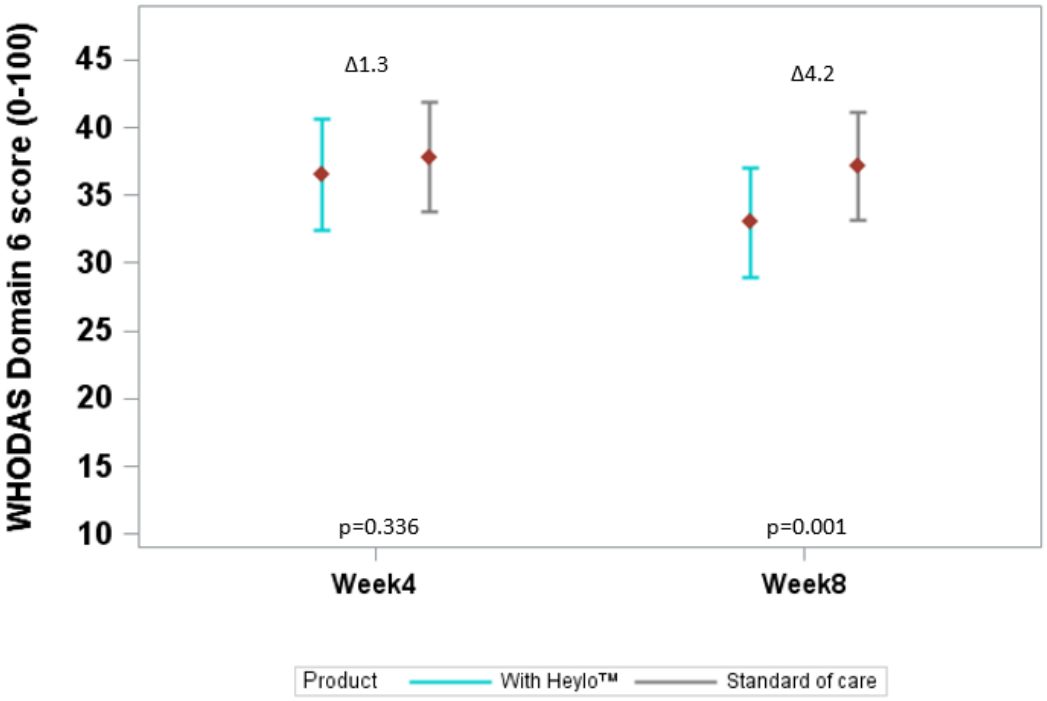


Figure 9 WHODAS 2.0 domain 6 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

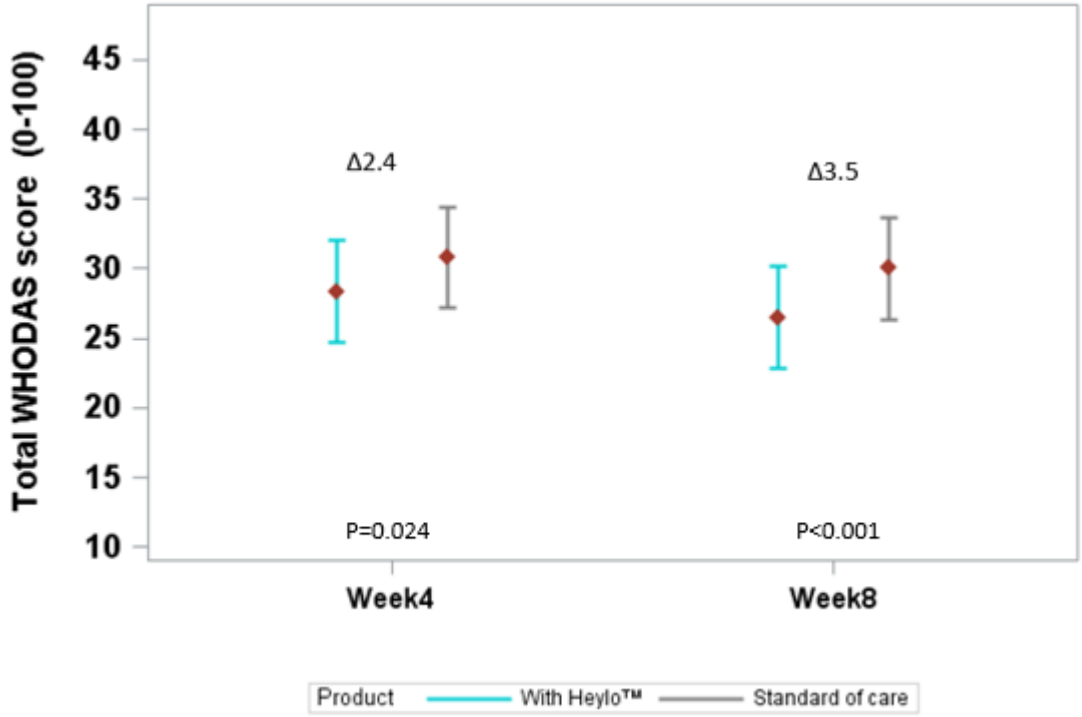


Figure 10 Total WHODAS 2.0 scores during the investigation. P-values, LS means and error bars representing the 95% confidence intervals of the LS means are displayed.

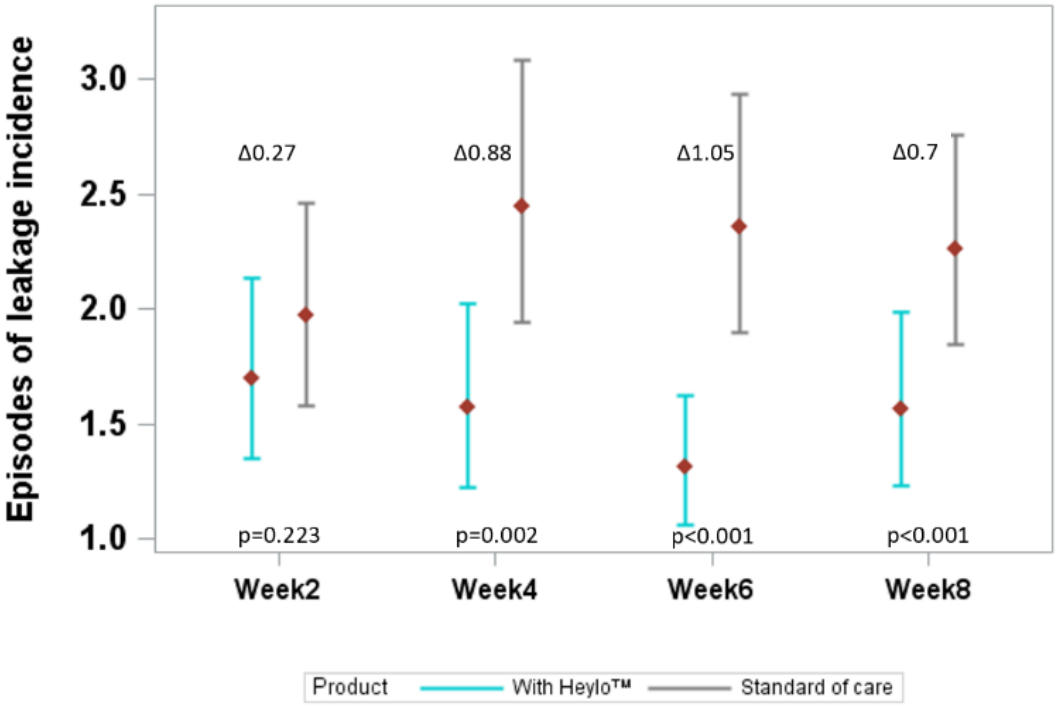


Figure 11 Episodes of leakage incidence outside baseplate onto clothes per 2 weeks during the investigation. LS means including 95% confidence intervals and P-values for episodes of leakage incidence outside baseplate onto clothes per 2 weeks during the investigation for Heylo™ and standard of care are displayed.

12.17 Detailed description of imputation methods

Multiple Imputation (MI) assuming Missing at Random (MAR)

1. **Imputation of missing values of the endpoint in period 1**.
    a. First intermittent missing values from baseline to week 8 in period 1 were imputed within each sequence (Heylo™/SoC and SoC/Heylo™) using a Markov Chain Monte Carlo (MCMC) method, to obtain a monotone missing data pattern. 100 imputations were applied.
    b. For each of the 100 datasets with a monotone missing data pattern in period 1, missing data were imputed sequentially within each sequence.
        i. First a linear model was fitted with the first planned value in period 1 as response and the baseline value as covariate. The estimated parameters and their variances from this model were used to impute missing values at the first planned timepoint in period 1.
        ii. Second, a linear model was fitted with the second planned value in period 1 as response and the baseline value as well as the first value in period 1 as covariates. The estimated parameters and their variances from this model were used to impute missing values at the second planned timepoint in period 1.
        iii. This step was repeated sequentially, in each step adding an additional value as a covariate, until missing week 8 assessments in period 1 were imputed.

2. **Imputation of missing values of the endpoint in period 2**.
    a. In each of the 100 datasets intermittent missing values from baseline and during period 2 were imputed within each sequence using a Markov Chain Monte Carlo (MCMC) method, to obtain a monotone missing data pattern in period 2.
    b. For each of the 100 datasets with a monotone missing data pattern in period 2, missing data were imputed sequentially within each sequence.
        i. First a linear model was fitted with the first planned value in period 2 as response and the baseline value as covariate. The estimated parameters and their variances from this model were used to impute missing values at the first planned timepoint in period 2.
        ii. Second, a linear model was fitted with the second planned value in period 2 as response and the baseline value as well as the first value in period 2 as covariates. The estimated parameters and their variances from this model were used to impute missing values at the second planned timepoint in period 2.
        iii. This step was repeated sequentially, in each step adding an additional value as a covariate, until missing week 8 assessments in period 2 were imputed.

3. **Analysis of the data**
    a. Each of the 100 datasets now had either observed or imputed week 8 assessments for all randomized subjects in both periods. For each dataset, an analysis was performed using a Mixed Model with the week 8 assessment as the response variable, treatment (Heylo™/SoC) and period (1/2) as fixed effects and a random effect of subject. In another analysis, baseline and the interaction between baseline and period were also included as covariates.
    b. The estimates and standard errors from the 100 datasets were pooled to one estimate and associated standard error using Rubin's rule. From these pooled estimates the confidence interval for the week 8 treatment contrast and the associated p-value were calculated.

Reference based Multiple Imputation (MI)

1. **Imputation of missing values of the endpoint in period 1**.
    a. First intermittent missing values from baseline to week 8 in period 1 were imputed within each sequence using a Markov Chain Monte Carlo (MCMC) method, to obtain a monotone missing data pattern. 100 imputations were applied.
    b. For each of the 100 datasets with a monotone missing data pattern in period 1, missing data were imputed sequentially. *The models used to impute the data were fitted only based on data from subjects randomized to SoC in period 1.*

        i. First a linear model was fitted with the first planned value in period 1 as response and the baseline value as covariate. The estimated parameters and their variances from this model were used to impute missing values at the first planned timepoint in period 1 for subjects randomized to Heylo™ as well as for subjects randomized to SoC.
        ii. Second, a linear model was fitted with the second planned value in period 1 as response and the baseline value as well as the first value in period 1 as covariates. The estimated parameters and their variances from this model were used to impute missing values at the second planned timepoint in period 1 for subjects randomized to Heylo™ as well as for subjects randomized to SoC.
        iii. This step was repeated sequentially, in each step adding an additional value as a covariate, until missing week 8 assessments in period 1 were imputed.

2. **Imputation of missing values of the endpoint in period 2**.
    a. In each of the 100 datasets intermittent missing values from baseline and during period 2 were imputed within each sequence using a Markov Chain Monte Carlo (MCMC) method, to obtain a monotone missing data pattern in period 2.
    b. For each of the 100 datasets with a monotone missing data pattern, missing data were imputed sequentially. *The models used to impute the data were fitted only based on data from subjects randomized to SoC in period 2.*
        i. First a linear model was fitted with the first planned value in period 2 as response and the baseline value as covariate. The estimated parameters and their variances from this model were used to impute missing values at the first planned timepoint in period 2 for subjects randomized to Heylo™ as well as for subjects randomized to SoC.
        ii. Second, a linear model was fitted with the second planned value in period 2 as response and the baseline value as well as the first value in period 2 as covariates. The estimated parameters and their variances from this model were used to impute missing values at the second planned timepoint in period 2 for subjects randomized to Heylo™ as well as for subjects randomized to SoC.
        iii. This step was repeated sequentially, in each step adding an additional value as a covariate, until missing week 8 assessments in period 2 were imputed.

3. Analysis of the data

    a. Each of the 100 datasets now had either observed or imputed week 8 assessments for all randomized subjects in both periods. For each dataset, an analysis was performed using a Mixed Model with the week 8 assessment as the response variable, treatment (Heylo™/SoC) and period (1/2) as fixed effects and a random effect of subject. In another analysis baseline and the interaction between baseline and period were also included as covariates.
    b. The estimates and standard errors from the 100 datasets were pooled to one estimate and associated standard error using Rubin's rule. From these pooled estimates the confidence interval for the week 8 treatment contrast and the associated p-value were calculated.

Reason for signing: Approved	[redacted]
Reason for signing: Approved	
Reason for signing: Approved	
Reason for signing: Approved	
Reason for signing: Approved	
Reason for signing: Approved	
Reason for signing: Approved	

Signature Page for VV-0336580 v4.0

Signature Page for VV-0336580 v4.0

Reason for signing: Approved	[redacted]

Signature Page for VV-0336580 v4.0