

**Burden of endometriosis: a mixed methods study
on costs and quality of life of adult women in the
Netherlands diagnosed with endometriosis by
laparoscopy, magnetic resonance imaging scan or
ultrasound scan**

Master Thesis Health Economics

Scientific Paper: PLoS One

Liselotte Tjon Joe Gin

581601

Thesis supervisor: LGM Bremmers

Second reader Reading Committee: prof.dr. EKA van Doorslaer

Erasmus School of Economics - Erasmus University Rotterdam

Word count: 11,531

05-07-2021

Abstract

Background and aim: Endometriosis is a chronic disease affecting women in the reproductive stage of their lives and causes severe pain and symptoms. Mixed methods research investigating the burden of endometriosis in the Netherlands is scarce. The aim of this study is to examine the burden of illness and quality of life of adult women in the Netherlands clinically diagnosed with endometriosis by laparoscopy, magnetic resonance imaging scan or ultrasound scan.

Methods: A mixed methods study design was employed by means of a retrospective cohort study that consisted of a burden of illness study and two focus group sessions. Direct health care and non-health care costs (i.e. iMCQ), indirect costs (i.e. iPCQ), and health-related-, endometriosis-related- and mental health-related quality of life (i.e. EQ-5D-5L, EHP-30 and MHQoL respectively) were determined together with the main drivers of annual health care costs and total costs. Two focus group sessions examined the feelings on, experiences with and thoughts on the burden of endometriosis of seven women to complement the quantitative results. Main theme, subthemes and categories were elicited during these sessions.

Results: The burden of illness study elicited annual total costs of on average €28,772 (95% CI: €23,276-€34,269). Productivity loss costs (mean: €20,303; 95% CI: €15,389-€25,217) represented 70.6% of the total costs. Average quality of life results were as follows: EQ-5D-5L value (0.585, SD: 0.277), EHP-30 score (average score of all five domains: 59.60, SD: 14.79) and MHQoL value (0.568, SD: 0.318). Examples of main drivers of annual health care costs and total costs were diagnostic delay and age categories 40-49 and 50-59. The long struggle was identified as the main theme, consisting of five subthemes: uncertainty, impact on daily life, frustration, awareness and maternity and infertility.

Conclusion: Endometriosis was found to have an extensive burden of disease for women in the Netherlands. Further research is needed to stress the societal relevance, raise awareness of the disease, reduce the diagnostic delay and thus decreasing the burden of endometriosis.

Table of contents

ABSTRACT	2
INTRODUCTION	4
METHODS	6
ETHICAL STATEMENT	6
BURDEN OF ILLNESS STUDY	6
<i>Survey description and data collection</i>	6
<i>Sampling and power analysis</i>	8
<i>Perspective</i>	8
<i>Data valuation</i>	8
<i>Data cleaning and assumption testing</i>	9
<i>Analysis</i>	10
<i>Post hoc analyses</i>	12
FOCUS GROUP SESSIONS	12
<i>Design</i>	12
<i>Sample and participants</i>	12
<i>Data collection</i>	14
<i>Data analysis</i>	14
<i>Study rigor</i>	15
RESULTS	15
RESULTS BURDEN OF ILLNESS STUDY	15
<i>Average yearly costs and quality of life</i>	16
<i>Multiple linear regression analyses</i>	18
<i>Post hoc analyses</i>	19
RESULTS FOCUS GROUP SESSIONS.....	21
<i>The long struggle</i>	22
<i>Uncertainty and day-to-day life</i>	23
<i>Impact on daily life</i>	23
<i>Frustration</i>	25
<i>Awareness</i>	25
<i>Maternity and infertility</i>	26
DISCUSSION	27
COSTS	27
QUALITY OF LIFE	29
MULTIPLE REGRESSION ANALYSES	30
FOCUS GROUP SESSIONS	31
STRENGTHS AND LIMITATIONS	32
CONCLUSION	34
ACKNOWLEDGEMENT	34
REFERENCES	34
APPENDICES	41
APPENDIX 1. ASSUMPTIONS OF THE STUDY	41
APPENDIX 2. DO-FILE DATA CLEANING	43
APPENDIX 3. DO-FILE ASSUMPTION TESTING	73
APPENDIX 4. ASSUMPTION TESTING; DETAILS	76
APPENDIX 5. DO-FILE INFERENTIAL STATISTICS AND POST HOC ANALYSES	78
APPENDIX 6. TOPIC LIST FOCUS GROUP SESSIONS	80

Introduction

Endometriosis is a chronic disease where endometrial tissue that normally grows inside the uterus also grows outside the uterus. As a result of this endometrial tissue, organs in the pelvic area (e.g. bladder, colon and ovaries) may become inflamed [1] which commonly leads to severe dysmenorrhea, excessive pelvic pain, ovulation pain, menstrual cycle associated symptoms, chronic fatigue, deep dyspareunia and infertility [2]. These symptoms are often attributed to gastrointestinal or gynecological conditions [3]. Five to fifteen percent of all women in the reproductive stage of their lives are affected by this disease. This amounts to approximately 400.000 women in the Netherlands [4]. However, a majority of these women with endometriosis are never diagnosed with the disease and a late diagnosis with a mean between five and ten years is common [5,6]. Median time between symptoms and diagnosis in the Netherlands is more than seven years [7]. This delay in diagnosis may be a consequence of patients', primary care practitioners' and specialists' low health literacy concerning endometriosis [5]. These aforementioned factors and symptoms contribute to a heavy burden on diagnosed women consisting of physical pain (i.e. quality of life), extensive medical treatments (i.e. medical costs), mental health problems (i.e. quality of life), productivity loss (i.e. economic burden), and alternative non-health care costs [1,8]. It is important to include all of these costs and effects of endometriosis in determining the true societal burden of illness [9,10].

A powerful practice to thoroughly examine the societal burden of endometriosis and to examine the burden of chronic diseases is a mixed methods study [11]. This is necessary to stress the need for early diagnosis and to restrict the physical (e.g. infertility/pain) and emotional damage to women [6,12,13]. A mixed methods study is preferred, since endometriosis is a complex illness and its burden differs substantially between women. These differences are for example present in pain severity, time between symptoms and diagnosis, endometriosis-related symptoms and prognosis [9]. Here, qualitative research complements the abstract results of a burden of illness study by examining a comprehensive understanding of the experiences, feelings and thoughts behind the unilateral costs side of endometriosis. Since endometriosis is a complex illness that seems to be not well known yet at societal level, focus groups may give a more in-depth insight in this true burden [14]. Even though societal impact seems to be high, little attention is given to endometriosis compared to other high-cost- and high-burden-diseases (e.g. rheumatoid arthritis and diabetes mellitus) [8,15]. Eventually, achieving societal awareness of endometriosis is the main aim of this research.

To the best of the researcher's knowledge, no mixed methods studies on the burden of endometriosis in the Netherlands can be found. An example of a mixed methods burden of illness study executed outside the Netherlands is Moradi et al. (2019) [16]. They executed a mixed methods study in which a clear correlation over time was found between endometriosis and the impact on different aspects of life amongst women with self-reported endometriosis in Australia. However, their quantitative section was focused on the validation of the qualitative questionnaire of the long-term impact of endometriosis and not on examining associated costs [16]. Individual burden of illness studies and qualitative studies representing focus group sessions are more common [15,17].

For the burden of endometriosis, a number of studies have been executed to understand the true effect of the illness. Examples of performed studies are Klein et al. (2014) [6] in Belgium, Soliman et al. (2019) [18] in the United States of America (USA) and Simoens et al. (2012) [15] in nine European countries (including the Netherlands) and the USA. Simoens et al. (2012) [15] is the only burden of disease study investigating the burden of endometriosis in the Netherlands. Its study population consisted of women clinically diagnosed with endometriosis and having at least one visit in a specialized clinic. For this sample, average total costs of endometriosis were elicited to be €9,579 per year (95% CI: €8,559-€10,599). Average annual productivity loss costs were equal to €6,298 (95% CI: €5,826-€6,771) and annual quality of life was estimated at an average of 0.809 Quality Adjusted Life Years (QALYs), where zero represents death and one represents perfect health [15]. These costs and quality of life were substantially affected by a delay in diagnosis. However, some concerns exist regarding the external validity of the results due to the sampling and recruitment methods employed. Moreover, Simoens et al (2012) [15] presented their results as an average over all 10 countries without indicating specific results for the Netherlands. Additionally, their study lacked subsequent qualitative research to complement and strengthen their findings.

For qualitative research, numerous studies have been executed by means of a solely qualitative method [17,19-23]. In general, these studies emphasized the negative influence of endometriosis on several domains in life. Examples of these domains were social life, education, work, physical ability, everyday activities, mental health, planning, having children and intimate relationships. A reduction in the negative consequences of the disease could be achieved by creating a more comprehensive understanding of endometriosis [20,22,23]. For example, in Ballweg (1997) [19] and Cox et al. (2003) [21] focus groups and interviews were executed to examine the effect of endometriosis on women's daily lives. The main themes elicited by these studies were the struggles, experiences with health care professionals and feeling misunderstood. Furthermore, adaptations in life and a lack of support were frequently mentioned themes [19,21]. However, no study exists on determining the experiences and feelings in daily life of women in the Netherlands with endometriosis, representing part of the burden of the illness.

To address the aforementioned concerns, this study will incorporate a mixed methods study to address the complexity of the disorder and strengthen the literature base. Hence, this study aims to investigate the true burden of illness (i.e. costs and quality of life) on a societal perspective for women diagnosed with endometriosis in the Netherlands. The following main question will be examined: *What is the burden of illness for adult women (aged 18 and above) in the Netherlands diagnosed with endometriosis by laparoscopy, magnetic resonance imaging (MRI) scan or ultrasound scan?* This question will be answered by means of a mixed methods explanatory sequential design study where a burden of illness study is followed by two focus group sessions. First, this study aims at determining the health care and non-health care costs by means of eliciting the descriptive statistics of these costs. Second, this study aims at determining the quality of life of women diagnosed with endometriosis also by means of descriptive statistical analyses. The third aim is to determine the main cost drivers of the burden of endometriosis in terms of health care costs and total costs. This will be done by means of multiple linear regression analyses. For both dependent variables, the following hypothesis will be tested:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$$

$$H_1: \beta_1 \neq 0 \text{ or } \beta_2 \neq 0 \text{ or } \beta_3 \neq 0 \text{ or } \beta_4 \neq 0 \text{ or } \beta_5 \neq 0 \text{ or } \beta_6 \neq 0 \text{ or } \beta_7 \neq 0$$

Lastly, this research aims at determining the experiences, feelings and thoughts accompanying daily lives of women diagnosed with endometriosis by the execution of two focus group sessions.

Methods

A mixed methods study was executed by the implementation of an explanatory sequential design. First, a retrospective prevalence-based burden of illness study was executed. A modified survey was created and published using Qualtrics software, version January-June 2021 [24]. Data was analyzed afterwards in Stata Release 16 [25]. The study was followed by two focus groups designed to establish an in-depth insight in the true burden of endometriosis, eliciting experiences, feelings and thoughts accompanying daily lives of women with endometriosis.

Ethical statement

The complete study was approved by the Research Ethics Review Committee of the Erasmus School of Health Policy and Management Rotterdam (approved February 1st 2021). All participants provided informed consent prior to conducting the survey and participating in the focus group session. Additionally, all participants of the focus group sessions provided oral consent prior to the start of the sessions.

Burden of illness study

The initial part of the research consisted of a burden of illness study where a survey was used to collect data. The checklist for a cost-of-illness (COI) study was used for reporting guidelines of this research [26].

Survey description and data collection

For the survey, an introductory section, the iMCQ [27], the iPCQ [28], the EQ-5D-5L [29], the EHP-30 [30], the MHQoL [31], a concluding section and a qualitative research recruitment question were included. In the introduction of the questionnaire, the objective of the study was introduced to the patients. Some introductory questions reflected the inclusion criteria (e.g. diagnosis, being an adult) and baseline characteristics (e.g. age, weight, area endometriosis, year of diagnosis). Consequently, five existing questionnaires (i.e. iMCQ [27], iPCQ [28], EQ-5D-5L [29], EHP-30 [30], MHQoL [31]) represented the main data collection on the health care costs, productivity costs and quality of life.

The generic iMCQ [27] measures self-reported medical consumption. It is a consistent valid and standardized questionnaire in the Netherlands to collect data on medical consumption as a result of physical or mental health problems [27]. The iPCQ [28] is a generic instrument to collect data on productivity loss focused on paid and unpaid work with a recall period of four

weeks. It is a consistent, valid and standardized questionnaire in the Netherlands that measures this productivity loss due to physical or psychological problems [28,32]. The EQ-5D-5L [33] is the five-dimensional EuroQol questionnaire that measure current health-related quality of life. Namely mobility, selfcare, daily activities, pain and discomfort and anxiety and depression. Each dimension has five answer levels ranging from one (i.e. the best option) to five (i.e. the worst option). Value sets representing utility values for different EQ-5D-5L health states were obtained from the national tariff for the five-level dimension of the Netherlands [33]. The EQ-5D-5L additionally contains a VAS-score question where 0 is the worst imaginable current health and 100 the best imaginable current health [29]. The EQ-5D-5L was found to be a reliable and valid estimate of health-related quality of life in the Netherlands [33]. The 30 questioned EHP-30 [30] examines the health-related quality of life for endometriosis patients. It contains five answer levels (i.e. never, rarely, sometimes, often, always) and is divided into five dimensions each containing a different amount of questions. These five dimensions are pain, control and powerlessness, social support, emotional well-being and self-image. Currently, no utility tariffs for the EHP-30 exist. However, for each dimension a score ranging from 0 to 100 is calculated accounting for the numbers of questions included in the dimension. Where 0 represents the best and 100 the worst imaginable endometriosis-related quality of life [30]. The EHP-30 was found to be a reliable and validated questionnaire measuring endometriosis-related quality of life in the Netherlands [34]. The MHQoL [31] is a validated standardized questionnaire in the Netherlands that measures the mental health-related quality of life for people with mental health problems [35]. It consists of seven dimensions that measure mental health-related quality of life of patients. The dimensions are self-image, independence, mood, relationships, daily activities, physical health and future. Each dimension has four answer levels, from one being very optimistic and not facing any problems until four being very pessimistic and facing many problems [31]. The MHQoL index score ranges from 0 to 21 and a higher score indicates better mental health-related quality of life [31]. Preliminary utility tariffs attached to the MHQoL scores were obtained by personal communication. The MHQoL [31] additionally contains a psychological well-being scale where 0 represents the worst and 10 the best imaginable psychological well-being. Since no utility values were present for the EHP-30, only the predicted index values for the EQ-5D-5L and MHQoL were used in the regression analysis for the purpose of measuring quality of life. EHP-30 scores, the EQ-5D-5L VAS-scores, the MHQoL psychological well-being scale and the MHQoL index scores were used for the baseline characteristics.

For the composition of the survey, several adjustments were made to the included existing questionnaires. This was done after extensive discussion between the researcher, a specialized gynecologist of Haaglanden Medical Center located in The Hague, and a health economists of the Erasmus University of Rotterdam. The adjustments aim at adapting the questionnaire for endometriosis specifically and the current circumstances of COVID-19. The iMCQ [27] was adapted to endometriosis-related health care use. Some questions without a direct relationship to endometriosis were removed. Moreover, the recall period was extended from three months to a more representative period of six months. This was done to account for the current COVID-19 circumstances where surgeries and treatments have been postponed [36]. The iPCQ [28], EQ-5D-5L [29] and MHQoL [31] were added without modifications. Lastly,

only the core questions of the EHP-30 [30] and not the full questionnaire on endometriosis-related quality of life were used for length-of-survey purposes.

For checking the validity of the questionnaire, a pilot study was executed after completing the composition and adaptation of the survey. Final adjustments were made after analyzing the feedback of three women recruited by the Endometriosis Foundation of the Netherlands. Consequently, the questionnaire was published online at the end of February 2021 and closed at the end of March 2021 representing a data collection period of one month.

Sampling and power analysis

Relevant patient inclusion criteria for the questionnaire are women in the Netherlands aged 18 or above diagnosed with endometriosis by laparoscopy, MRI scan or ultrasound scan. This clinical diagnosis was chosen as an inclusion criterion since a diagnosis can only be assured with these means. Without a laparoscopy, an MRI scan or an ultrasound scan physicians cannot indicate the presence of endometriosis with certainty [37]. Additionally, the grade of endometriosis should be familiar to the participant and the participant should have given written consent to be included in the study. The grade of endometriosis is a grading system used by Dutch physicians to determine the severity of the illness, with grade one and two being superficial and mild, while grade three and four are indicated as deep invasive endometriosis [37]. The clinical diagnosis and grade are self-reported by participants which we assume to be indicated truthfully. The Endometriosis Foundation of the Netherlands was asked to share the link to the online survey to their social media platforms to recruit participants for the survey.

An a priori power analysis was conducted in G*Power 3.1.9.7 [38] in order to calculate the required sample size for the multiple linear regression analysis for both outcome variables; annual health care costs and annual total costs. The power analysis was done with the inclusion of seven independent variables, the significance level of 5%, a one tailed test, the effect size of 0.15 and power of 0.80. The total sample size for a meaningful effect was calculated to be 43.

Perspective

To examine the overall impact of endometriosis, a societal perspective was applied. This perspective incorporates all relevant costs across all relevant sectors, independent of who pays for it [39]. The direct costs of health care use, like medication and physician visits, the impact on productivity loss and associated costs, family costs (e.g. support) and travel costs were incorporated. Additionally, health-related quality of life, endometriosis-related quality of life and mental health-related quality of life were considered.

Data valuation

Total costs in the analysis consist of direct health care costs (i.e. directly related to health care), direct non-health care costs and indirect costs. For extensive definitions of these cost variables see Table 1. For indirect cost measures, the friction cost method was used, incorporating vacancies and the friction period [40]. All prices in the Dutch manual for costing studies [40] were Euro 2014 reference prices and were discounted to 2020 prices with the annual Dutch CPI change from 2014 to 2020 [40,41]. For the calculation of the medication costs where use, frequency and/or dose were unknown in the data, the conservative mechanism was applied. Here, the unknown elements were set equal to zero. All medical care use of the

iMCQ [27] was multiplied by two to transform these costs from six months to a year. Productivity costs of the iPCQ [28] concerning short paid work absenteeism, paid work presenteeism and unpaid work presenteeism in the past four weeks were multiplied by 13 to derive annual productivity costs. Long absenteeism costs, indicating more than four weeks, were calculated according to the friction cost method in the Dutch manual for costing studies [40].

Table 1. Definitions cost variables

Resource	Variable	Source of cost	Measurement
Direct health care costs		All self-reported by participants. Reference price values used from the Dutch manual for costing studies [40] and average prices on the internet. Hormonal treatment dosages and prices used from medicijnkosten.nl [42] and Farmacotherapeutisch Kompas [43].	Measured by iMCQ [27] and over the past six months.
<i>Visits to other health care providers</i>	Type and number		
<i>Physician visits</i>	Type and number		
<i>Medication</i>	Type, dose, frequency, number of days		
<i>Treatments (outpatient)</i>	Type and number		
<i>Hospitalization (inpatient)</i>	Number of days and frequency		
<i>Ambulance care</i>	Number		
Direct non-health care costs		All self-reported by participants. Reference price values used from the Dutch manual for costing studies [40].	Measured by iMCQ [27] and over the past six months.
<i>Travel costs</i>	Way of transport and times		
<i>Support household activities</i>	Type and number		
Indirect costs		All self-reported by participants. Friction cost method and reference prices per hour from the Dutch manual for costing studies [40].	Measured by iPCQ [28] and over past four weeks.
<i>Productivity loss work</i>	Number of days, number of hours; absenteeism, presenteeism; paid and unpaid work		

Data cleaning and assumption testing

During data cleaning in Stata Release 16 [25], all distinct irregularities in the dataset were removed. Examples of processes of the data cleaning are renaming variables, dropping variables that were not useful, creating new variables like BMI and dummies for comorbidities and area, labeling and defining variables, ordering variables, implementing the inclusion criteria, adding codes for utility application and dropping observations which represented irregularities. Several assumptions were made in order to clean the data and accurately perform data analysis. These assumptions and possible consequences are stated in Appendix 1. The do-file of the data cleaning and valuation can be found in Appendix 2. Subsequently, forced entry was used to decide upon the regressors to include. The choice of predictors was based on previous research [15] and on thoroughly discussion on theoretical relevance with gynecologists and endometriosis specialists of Haaglanden Medical Center located in The Hague.

The assumptions of a multiple linear regression analysis were tested in order to be able to fulfill the generalizability of the model (do-file in Appendix 3). For the assumption testing, OLS assumptions were tested in Stata Release 16 [25]. Non-zero variance in predictors and no

perfect multicollinearity between two or more predictors should be present. The model should be well specified (i.e. predictors being uncorrelated with other variables), there should be no strong evidence of heteroskedasticity and errors should be normally distributed. Independence should be assumed and a linear relationship between predictors and outcome variables should exist [44].

The study population was extremely heterogeneous and no concerns existed on outliers in the population sample. Consequently, no concerning results were found during the assumption testing except for the multicollinearity and omitted variable assumption. A VIF-value of 3.49 was found when testing for multicollinearity for both outcome variables, thus not completely fulfilling the no perfect multicollinearity assumption. This may be the result of the extensively high VIF-values for the grade of endometriosis, ranging from 5.88 to 11.31. This may have been caused by the fact that the population was skewed towards grade four of endometriosis and grade one and two were underrepresented in the sample. There were some concerns on the specification of the model for annual total costs (i.e. linktest and ovtest). This was probably the case since the forced entry method was used. Even though this may be a problem, due to the strong theoretical framework (i.e. existing literature and extensive discussion with endometriosis specialists) the used method was assumed to be valid. For all tests, the grade of endometriosis seem to be the most “violating” regressor for the assumption testing. A detailed description and the results of the assumption testing can be found in Appendix 4. Throughout the study, a significance level of 5% was used.

Analysis

For the descriptive statistics, baseline characteristics of the sample were elicited. These characteristics contain information on different demographics of the sample population. The demographics consist of age (18-29 year/30-39 years/40-49 years/50-59 years), height (meters), weight (kilograms), BMI (kilograms/m²), marital status (single/married/single and living with a partner/divorced or separated/widowed), highest attained educational level (no education/primary school/domestic science school/secondary school/vocational education and training (VET)/higher professional education/research-oriented education), occupation (student/employee/self-employed/housewife/not employed/unable to work/(early) retired/other), degree of incapacity to work (percentage) if participant indicated to be unable to work and absenteeism (average number of days of part of sample with a paid job). Clinical characteristics like area of endometriosis (peritoneum/bladder/colon/ovaries/adenomyosis/vagina/urethra/ligaments/thoracic/other), diagnostic delay (years) and comorbidities (e.g. infertility/migraine/depression/backpain/fatigue/other/none) were reported as well. Means of these descriptive statistics were computed accompanied by its percentage (%), standard deviation (SD) and minimum and maximum.

Furthermore, costs were calculated by multiplying the visits and frequencies with the discounted reference price (i.e. discounted by CPI [41]) to get Euro 2020 costs for every woman individually. Additionally, average quality of life in terms of health, endometriosis and mental health were included. Mean costs and quality of life were computed accompanied by its standard deviation, interval of minimum and maximum and the 95% confidence interval. For

the descriptive statistics, direct health care costs, direct non-health care costs, indirect costs, total costs and quality of life measures were presented.

For the inferential statistics, two multiple linear regression analyses (i.e. OLS) were done to find the main cost drivers of yearly health care costs and yearly total costs. A log-transformation of the two outcome variables accounted for non-normality in these variables. The independent variables in both regressions were health-related quality of life, mental health-related quality of life, diagnostic delay, grade of endometriosis, BMI, age and the presence of infertility. An outline of these underlying predictors is stated in Table 2.

Table 2. Dependent and independent variables of the regression analyses

Dependent variable	Definition variable	Type of variable	Levels of measurement	Reference
Yearly health care costs	Consists of physician visits, other health care provider visits, medication, treatments, hospitalization and ambulance care.	Continuous - discrete	N/A	Simoens et al. (2012) [15]
Yearly total costs	Consists of yearly health care costs, travel costs, costs for support and costs of productivity loss.	Continuous – discrete	N/A	Simoens et al. (2012) [15]
Independent variable	Definition variable	Type of variable	Levels of measurement	Reference
Health-related quality of life	EQ-5D-5L [29] index values with utility tariffs attached [33]. 0 depicts death and 1 depicts perfect health.	Continuous – Discrete (EQ-5D-5L [29] was measured as categorical-ordinal)	N/A	Simoens et al. (2012) [15]
Mental health-related quality of life	MHQoL [31] index values with preliminary utility tariffs attached obtained by personal communication. 0 depicts death and 1 depicts perfect health.	Continuous – Discrete (MHQoL [31] was measured as categorical-ordinal)	N/A	Extensive discussion with research team, specialized gynecologist and other supervisors of Erasmus University Rotterdam.
Diagnostic delay	Time between first symptoms and diagnosis in years.	Continuous - Discrete	N/A	Simoens et al. (2012) [15]
Grade of endometriosis	Different grades of endometriosis.	Categorical - Ordinal	Grade 1, grade 2, grade 3, grade 4	Simoens et al. (2012) [15]
Body Mass Index (BMI)	Bodyweight in kg divided by height in m ² (kg/m ²).	Continuous - Discrete	N/A	Simoens et al. (2012) [15]
Age	Age in years of respondents divided into 4 categories.	Categorical - Ordinal	18-29 years, 30-39 years, 40-49 years, 50-59 years	Simoens et al. (2012) [15]
Presence of infertility	Whether or not participant deals with infertility.	Categorical - Binary	1 = yes, 0 = no	Simoens et al. (2012) [15]

The formula used to calculate the magnitude of the effect in a log-linear model is stated in formula 1:

$$\% \Delta y = 100 * (e^{\beta_i} - 1) \quad (1)$$

To test the overall significance of the model, an F-test was done in Stata Release 16 [25]. The power of the model was judged by the goodness of fit in R^2 and adjusted R^2 which represent the cross validation of the study.

Post hoc analyses

Two sensitivity analyses and two ANOVA tests were executed as post hoc analyses. For the first sensitivity analysis, the variable representing the grade of endometriosis was excluded from the regression. Namely, because the variable seems to be skewed towards higher grades which may impact the results of the regression analysis. For the second sensitivity analysis, mental health-related quality of life was excluded as a quality of life measure in the regression. Both quality of life measures (i.e. EQ-5D-5L and MHQoL) were included due to the major theoretical effect of endometriosis on mental health-related quality of life and not just health-related quality of life. However, both measures have some overlap in for example daily activities and the mental health-related aspect in both measures. This analysis was done in order to see the effect on the regression results when the collinearity between the two measures that may be present was eliminated. For the ANOVA tests, different grades of endometriosis and different age categories were compared in outcomes. According to the properties of the dataset, it made sense to execute an ANOVA test for sub group analysis. The do-file of the inferential statistics and post hoc analyses can be found in Appendix 5.

Focus group sessions

Design

A qualitative design was used to explore the feelings, experiences and thoughts behind the outcomes of the survey filled in by women in the Netherlands diagnosed with endometriosis. For reporting the methods and the results of the focus group sessions, the COREQ guidelines were applied [45].

Sample and participants

For sampling the focus group participants, respondents were able to indicate whether or not they wanted to participate in the focus group sessions at the end of the survey by leaving their email address. Participants were thus selected by a form of snowball sampling in which women themselves could indicate if they wanted to participate or not. These women received an email including information on the focus groups and with the question whether they still wanted to participate or not accompanied by an informed consent form. The first seven women responding to this email were chosen as the study sample. Since all participating women finished the survey, the sample consisted of adult women in the Netherlands clinically diagnosed with endometriosis. After conducting the burden of illness study, homogeneity was

observed in the sample concerning age and grade of endometriosis of the participants. Due to this fact, two sessions were assumed to be sufficient to reach saturation in the available sample.

Demographics and clinical characteristics of the participants can be found in Table 3. The average age of participants was 36.7 years (SD: 6.50). Four participants were employed (57.1%) being a teacher, psychotherapist, a security guard and a nurse. One woman was unable to work (14.3%) due to endometriosis and one woman was unemployed due to endometriosis (14.3%). Occupation of one participant was unknown (14.3%). Most common areas of growing endometrial tissue in the sample population were the colon (85.7%) and the bladder (71.4%). Adenomyosis, which is endometrial tissue in the uterine wall, was most often represented in the sample population (85.7%). Examples of common comorbidities in the sample were difficulties with the colon (71.4%), depression (28.6%), chronic cystitis (28.6%), and difficulties with the bladder (28.6%).

Table 3. Baseline characteristics focus group sessions (n=7 women)

Characteristic	Mean (percent)
Age (years)	
18-29 years	1 (14.29%)
30-39 years	3 (42.86%)
40-49 years	3 (42.86%)
Occupation	
Employed	4 (57.14%)
Not employed	1 (14.29%)
Unable to work	1 (14.29%)
Unknown	1 (14.29%)
Marital status	
Married	4 (57.14%)
Relationship	2 (28.57%)
Single	1 (14.29%)
Reported area*	
Adenomyosis	6 (85.71%)
Colon	6 (85.71%)
Bladder	5 (71.43%)
Ovaries	2 (28.57%)
Fallopian tube	2 (28.57%)
Other organs	2 (28.57%)
Vagina	1 (14.29%)
Reported comorbidities*	
Difficulties with colon	5 (71.43%)
Depression	2 (28.57%)
Chronic cystitis	2 (28.57%)
Difficulties passing bladder	2 (28.57%)
Coeliac disease	1 (14.29%)
Mastocytosis	1 (14.29%)
Migraine	1 (14.29%)
Pulmonary embolism	1 (14.29%)
Thrombosis	1 (14.29%)

Note: *These were endometriosis areas and comorbidities that were explicitly reported by the respondents during the focus group sessions. Hence, this may not be a comprehensive list. Also, comorbidities were only registered when literally mentioned during the sessions. True number of participants with these comorbidities may be higher because not all were literally quoted during sessions.

Data collection

In total, seven women participated in the two focus group sessions (n=3 and n=4). Both sessions were approximately two hours and were in the Dutch language. The focus groups took place during May 2021. Due to COVID-19 circumstances, sessions were online via a protected connection on Zoom version 5 [46]. Audio recordings were used to collect the data.

The first author is a female master's degree student, BS Utrecht University School of Economics, and she conducted the focus groups, being present as the moderator of the sessions. No personal connection, relationship with participants or association of the author with the subject endometriosis was present before the start of the study. The only knowledge of the subject was present due to reading scientific papers and investigating into the subject using objective sources. The author recognized that designing the research set-up could lead to confirmation bias, representing behavior that leads towards findings needed for confirmation of the study. For this reason, continuous objective reflection during the analysis and discussion within the research group (of which Leonie Bremmers, PhD Candidate Erasmus School of Health Policy & Management is a member) took place [47]. A bracketing approach was used by the researcher during the focus group sessions. This consisted of opening up to the experiences and feelings of participants with the means of not applying the empirical and analytical mindset. In all this, the author was constantly aware of the main goal of empathic neutrality [47,48].

Before the start of the focus groups, the aim and background of the sessions were explained by the moderator. Notes were made during and directly after the two sessions to reflect on later in the process. A semi-structured format was followed. Topics for the topic list were identified after completing the burden of illness analysis and were additionally based on existing literature [17]. In Appendix 6 the topic list and the prompts for open-ended questions of the focus group sessions can be found.

Data analysis

After the focus group sessions, the recorded audio was transcribed verbatim and analyzed. A deductive analysis in ATLAS.ti 9 Mac [49] was executed for analyzing and interpreting the focus groups. Thus, the basis of the focus group sessions was built upon existing literature [17] and the results of the burden of disease study. The main theme and subthemes were derived from the results of the focus group data. Categories, being part of the subthemes, were examined via co-occurrence in ATLAS.ti 9 Mac [49]. Quotations were translated into English by the researcher. During the whole process of constructing the topic list, creating codes, analyzing the data and translating into English, critical reviews were present within the research group. Details of the analysis can be found in Table 4.

Table 4. Deductive analysis in ATLAS.ti 9 Mac [49]

Steps of deductive analysis	
1.	Recordings are listened to for multiple times to generate the transcripts.
2.	Transcripts are read thoroughly several times and are given codes.
3.	All codes were converted into themes, subthemes and categories by comparing them.
4.	Transcripts are re-read and recoding took place into the main themes, subthemes and categories.
5.	Themes, subthemes and categories were reviewed and a complete understanding of the thoughts, experiences and feelings of diagnosed women with endometriosis is formed.
6.	Codes and themes, subthemes and categories were checked and discussed within the research group.

Study rigor

In depth discussions within the research group took place to enhance consistency of the transcripts and data analysis. Additionally, coding trees, final themes, subthemes and categories were discussed extensively within the same research group for the same reason.

Results

Results burden of illness study

A sample of 215 respondents was collected from the survey. After cleaning the data, a sample size of 107 respondents was left due to errors in incomplete responses and applying the inclusion criteria. In Table 5 demographics and clinical characteristics of participants are demonstrated. The majority of the sample was aged between 30-39 years old (43.0%). Another important characteristic of this study was the mean diagnostic delay of 12.17 years (SD: 7.43). Common areas of endometriosis were the ovaries (72.9%), colon (69.2%) and adenomyosis (52.3%). Major elicited comorbidities were fatigue (28.0%), backpain (25.2%), infertility (25.2%) and other comorbidities (46.7%). Examples of other comorbidities mentioned in the sample were asthma and polycystic ovarium syndrome (PCOS). Average days of absenteeism amounted to 40.28 (SD: 177.77), which is more than the indicated 28 days (four weeks) of the questionnaire.

Table 5. Baseline characteristics survey (n=107 women)

Characteristic	Mean (percent/SD, minimum-maximum)
Age (years)	
18-29 years	28 (26.17%)
30-39 years	46 (42.99%)
40-49 years	31 (28.97%)
50-59 years	2 (1.87%)
Height (m)	169.87 (6.69, 153-186)
Weight (kg)	73.69 (16.70, 51-125)
BMI	25.52 (5.40, 17.71-41.77)
Marital status	
Single	35 (32.71%)
Married	43 (40.19%)
Single and living with partner	27 (25.23%)
Divorced/separated	2 (1.87%)
Widowed	0 (0%)

Highest attained educational level*	
<i>No education</i>	0 (0%)
<i>Primary school</i>	0 (0%)
<i>Domestic science school</i>	0 (0%)
<i>Secondary school</i>	10 (9.34%)
<i>Vocational education and training (VET)</i>	32 (29.91%)
<i>Higher professional education</i>	43 (40.19%)
<i>Research-oriented education</i>	22 (20.56%)
Occupation	
<i>Student</i>	4 (3.74%)
<i>Employed</i>	62 (57.94%)
<i>Self-employed</i>	13 (12.15%)
<i>Housewife</i>	1 (0.93%)
<i>Not employed</i>	2 (1.87%)
<i>Unable to work</i>	17 (15.89%)
<i>(Early) Retired</i>	0 (0%)
<i>Other</i>	8 (7.48%)
Degree of inability to work (%; n=15)**	87.47 (21.54, 34-100)
Grade endometriosis	
<i>Grade 1</i>	2 (1.87%)
<i>Grade 2</i>	10 (9.35%)
<i>Grade 3</i>	17 (15.89%)
<i>Grade 4</i>	78 (72.90%)
Diagnostic delay (years)	12.17 (7.43, 0-32)
Area endometriosis	
<i>Ovaries</i>	78 (72.90%)
<i>Colon</i>	74 (69.16%)
<i>Adenomyosis</i>	56 (52.34%)
<i>Peritoneum</i>	52 (48.60%)
<i>Bladder</i>	47 (43.93%)
<i>Urethra</i>	23 (21.50%)
<i>Vagina</i>	20 (18.69%)
<i>Ligaments</i>	16 (14.95%)
<i>Thoracic</i>	8 (7.48%)
<i>Other</i>	14 (13.08%)
Comorbidities	
<i>Fatigue</i>	30 (28.04%)
<i>Backpain</i>	27 (25.23%)
<i>Infertility</i>	27 (25.23%)
<i>Migraine</i>	20 (18.69%)
<i>Depression</i>	14 (13.08%)
<i>Other</i>	50 (46.73%)
<i>None</i>	26 (24.30%)
Absenteeism (days; n=81)***	40.28 (117.77, 0-845)

Note: *Educational levels were transformed from the Dutch system towards an international classification [50].

**n=15. 15 women indicated to be unable to work. The average degree of inability to work in percentage is stated here (standard deviation, minimum-maximum).

*** n=81. 81 women indicated to have a paid job. This is an indication of the average absent working days of these women.

Average yearly costs and quality of life

Yearly direct health care costs amounted to €5,323 (95% CI: €4,027-€6,619) where hospitalization (27.1%) and other visits (22.1%), like physical therapist visits, accounted for the biggest share of this. Yearly direct non-health care costs amounted to €3,146 (95% CI: €2,079-€4,214) per patient with its biggest share represented by support in household activities (90.2%). Substantial costs were elicited in average yearly productivity loss which amounted to €20,303 (95% CI: €15,389-€25,217) per patient and accounted for 70.6% of total costs. Total

costs amounted to €28,772 (95% CI: €23,276-€34,269) per patient per year. Results for costs are represented in Table 6.

Table 6. Annual costs of endometriosis per patient (n=107 women) (Euro 2020 (€))

Cost	Mean (SD)	Minimum, Maximum	95% Confidence Interval (CI)
Direct health care costs	5 323 (6 760)	71, 41 365	4 027, 6 619
<i>GP and GP-based nurse specialist</i>	189 (219)	0, 1 071	147, 231
<i>Other visits</i>	1 177 (1 439)	71, 9 831	902, 1 453
<i>Gynecologist visits</i>	504 (512)	0, 3 937	406, 602
<i>Fertility specialist visits</i>	155 (644)	0, 3 937	31, 278
<i>Other specialist visits (including ER doctor)</i>	460 (689)	0, 3 513	328, 592
<i>Hormonal therapy medication</i>	387 (607)	0, 1 723	270, 503
<i>Treatments</i>	988 (2 632)	0, 15 521	483, 1 492
<i>Hospitalization</i>	1 443 (4 766)	0, 30 887	530, 2 357
<i>Ambulance care</i>	21 (152)	0, 1 114	-8, 50
Direct non-health care costs – family costs	3 146 (5 570)	2, 40 922	2 079, 4 214
<i>Travel costs</i>	308 (808)	2, 7 282	154, 463
<i>Support household activities</i>	2 838 (5 499)	0, 40 880	1 784; 3 892
Total direct costs	8 469 (9 680)	73, 50 799	6 614, 10 325
Indirect costs			
<i>Productivity loss work</i>	20 303 (25 638)	0, 132 268	15 389, 25 217
Total costs	28 772 (28 679)	73, 138 511	23 276, 34 269

The average EQ-5D-5L index value amounted to 0.585 (SD: 0.277). The VAS-score was on average 55.33 (SD: 18.08) on a scale from 0 to 100. The average MHQoL index value equaled 0.568 (SD: 0.318) with a psychological well-being score of 6.23 (SD: 1.71) on a scale from 0 to 10. The average unweighted MHQoL index score was equal to 10.24 (SD: 2.14). EHP-30 scores ranged from 59.02 (SD: 18.80) to 65.20 (SD: 20.42) on the five different domains. Lowest average scores were elicited in support (mean: 59.02, SD: 18.80) and in self-image (mean: 59.56, SD: 20.05). The average score over the five domains amounted to 59.60 (SD: 14.79). The results for the quality of life measures can be found in Table 7.

Table 7. Quality of life outcomes (n=107 women)

Quality of life measure	Mean (SD)	Minimum, Maximum	95% CI
Health-related quality of life			
EQ-5D-5L index value (0-1)	.585 (.277)	-.137, 1	.53, .64
EQ-5D-5L VAS score (0-100)	55.33 (18.08)	10, 100	51.86, 58.79
Endometriosis-related quality of life			
EHP-30 value (0-100)*	59.60 (14.79)	20, 86.12	56.76, 62.43
<i>Pain</i>	60.39 (18.32)	20, 92.73	56.88, 63.90
<i>Control an powerlessness</i>	65.20 (20.42)	20, 100	61.29, 69.12
<i>Emotional wellbeing</i>	53.80 (15.18)	20, 96.67	50.89, 56.71
<i>Social support</i>	59.02 (18.80)	20, 95	55.42, 62.62
<i>Self-image</i>	59.56 (20.05)	20, 100	55.72, 63.41
Mental health-related quality of life			
MHQoL index score (0-21)	10.24 (2.14)	0, 14	9.83, 10.65
MHQoL index value (0-1)	.568 (.318)	-.687, 1	.507, .629
MHQoL psychological well-being (0-10)	6.23 (1.71)	2, 10	5.90, 6.56

Note: *Average value of the five domains

Multiple linear regression analyses

For the annual health care costs only diagnostic delay (95% CI: 0.01, 0.07; $p=0.013$) and age category 40-49 (95% CI: -1.73, -.35; $p=0.004$) and 50-59 (95% CI: -4.74, -1.25 ; $p=0.001$) seemed to be significant determinants of log yearly health care costs at a 5% significance level. An increase of one year in diagnostic delay correlates with an average increase in yearly health care costs of 4.2%, *ceteris paribus*. Also, being 40-49 years old is associated with 64.6% lower average yearly health care costs compared to being 18-29 years old, *ceteris paribus*. Being 50-59 years old is associated with an approximate 95.0% lower average yearly health care costs compared to being 18-29 years old, *ceteris paribus*.

For the annual total costs only the age category 50-59 (95% CI: -4.80, -1.62; $p=0.000$) was a significant determinant at a 5% significance level. Being 50-59 years old is associated with an approximate 96.0% lower average yearly health care costs compared to being 18-29 years old, *ceteris paribus*. All other regressors seemed to be insignificant at a 5% significance level. Results of the multiple linear regression analyses for annual health care costs and annual total costs are stated in Table 8. Formula 1 was used to calculate the magnitudes of the effects of the significant regressors.

Table 8. Results multivariate linear regression analysis for yearly health care costs and yearly total costs (n=107 women)

Independent variable	Yearly health care costs (log)			Yearly total costs (log)		
	Coefficient (SE)	p-value	95% CI	Coefficient (SE)	p-value	95% CI
Health-related quality of life	-.230 (.570)	0.687	-1.36, .90	-.668 (.519)	0.201	-1.70, .36
Mental health-related quality of life	-.250 (.489)	0.611	-1.22, .72	-.881 (.445)	0.051	-1.76, .00
Diagnostic delay (years)	.041 (.016)*	0.013	.01, .07	.022 (.015)	0.148	-.01, .05
Grade endometriosis						
2 (grade 2)	.875 (.922)	0.345	-.96, 2.71	-.406 (.839)	0.629	-2.07, 1.26
3 (grade 3)	.540 (.891)	0.546	-1.23, 2.31	-.378 (.811)	0.642	-1.99, 1.23
4 (grade 4)	.767 (.837)	0.362	-.89, 2.43	-.541 (.762)	0.479	-2.05, .97
BMI	.011 (.023)	0.634	-.04, .06	.024 (.021)	0.259	-.02, .07
Age categories						
2 (30-39 years)	-.281 (.298)	0.349	-.87, .31	-.247 (.271)	0.364	-.79, .29
3 (40-49 years)	-1.037 (.348)*	0.004	-1.73, -.35	-.363 (.317)	0.255	-.99, .27
4 (50-59 years)	-2.997 (.880)**	0.001	-4.74, -1.25	-3.213 (.801)**	0.000	-4.80, -1.62
Presence of infertility	.360 (.275)	0.194	-.19, .91	-.084 (.250)	0.738	-.58, .41
Constant	7.107 (1.087)	0.000	4.95, 9.26	10.571 (.989)	0.000	8.61, 12.53
R ²	.222	-	-	.274	-	-
Adjusted R ²	.132	-	-	.190	-	-
Overall F-test	F = 2.46 (p=0.009)			F = 3.26 (p<0.001)		

Note: * $p < 0.05$, ** $p < 0.01$

P-values of both models with the two outcome variables showed a lower value than the significance level of 5%. For log yearly health care costs $p=0.009$ ($F(11, 95)=2.46$) and R^2 is 0.222. For log yearly total costs $p<0.001$ ($F(11, 95)=3.26$) and R^2 is 0.274. We can reject H_0 and we may accept H_1 that this model fits the data better than a model without any regressors.

Alternatively, R^2 and adjusted R^2 were slightly higher for the log yearly total costs model compared to log yearly health care costs. However, both values were relatively low.

Post hoc analyses

For the two sensitivity analyses, the grade of endometriosis and the mental health-related quality of life were excluded independently. Grades of endometriosis were included for the sake of hypothetical relevance. Removing grade of endometriosis from the regression led to the results represented in Table 9. Only mental health-related quality of life became significant for annual total costs (95% CI: -1.79, -.05; $p=0.038$) at a 5% significance level. No other variables became significant and signs, magnitudes, standard errors and p-values only changed slightly. R^2 dropped somewhat compared to the main model (i.e. from 0.222 to 0.210 for yearly health care costs and from 0.274 to 0.268 for yearly total costs) which might be due to the fact that one variable was excluded. However, adjusted R^2 increased compared to the main model (i.e. from 0.132 to 0.145 for yearly health care costs and from 0.190 to 0.208 for yearly total costs). Overall goodness of fit and predictive power in the F-statistic seem to be improved.

Table 9. Results multivariate linear regression analysis with exclusion of grade of endometriosis (n=107 women)

Independent variable	Yearly health care costs (log)			Yearly total costs (log)		
	Coefficient (SE)	p-value	95% CI	Coefficient (SE)	p-value	95% CI
Health-related quality of life	-.288 (.562)	0.610	-1.40, .83	-.633 (.510)	0.217	-1.65, .38
Mental health-related quality of life	-.209 (.483)	0.666	-1.17, .75	-.920 (.438)*	0.038	-1.79, -.05
Diagnostic delay (years)	.041 (.016)*	0.012	.01, .07	.021 (.014)	0.151	-.01, .05
BMI	.009 (.023)	0.698	-.04, .05	.024 (.021)	0.256	-.02, .06
Age categories						
2 (30-39 years)	-.299 (.284)	0.296	-.86, .27	-.268 (.258)	0.302	-.78, .24
3 (40-49 years)	-1.056 (.337)*	0.002	-1.72, -.39	-.389 (.306)	0.206	-1.00, .22
4 (50-59 years)	-2.975 (.868)**	0.001	-4.70, -1.25	-3.267 (.787)**	0.000	-4.83, -1.70
Presence of infertility	.383 (.266)	0.153	-.14, .91	-.092 (.241)	0.705	-.57, .39
Constant	7.991 (.633)	0.000	6.65, 9.17	10.119 (.574)	0.000	8.98, 11.26
R^2	.210	-	-	.268	-	-
Adjusted R^2	.145	-	-	.208	-	-
Overall F-test	F = 3.25 ($p=0.003$)	-	-	F = 4.47 ($p<0.001$)	-	-

Note: * $p < 0.05$, ** $p < 0.01$

Mental health-related quality of life was included in the model to account for the predicted severe impact on women with endometriosis. Removing the mental health-related quality of life and only including the health-related quality of life, like in Simoens et al. (2012) [15], led to the results represented in Table 10. Health-related quality of life became significant (95% CI: -2.13, -.62; $p=0.000$) (and magnitude became substantially larger) for this variable only for the annual total costs regression at a 5% significance level. No other variables became significant and sign, magnitude, standard errors and p-values only changed slightly. However,

compared to the main model, overall goodness of fit and predictive power in the F-statistic increased slightly for log yearly health care costs. For log yearly total costs, goodness of fit and predictive power decreased compared to the main model.

Table 10. Results multivariate linear regression analysis with exclusion of mental health-related quality of life (n=107 women)

Independent variable	Yearly health care costs (log)			Yearly total costs (log)		
	Coefficient (SE)	p-value	95% CI	Coefficient (SE)	p-value	95% CI
Health-related quality of life	-.432 (.409)	0.294	-1.24, .38	-1.379 (.380)**	0.000	-2.13, -.62
Diagnostic delay (years)	.040 (.016)*	0.014	.01, .07	.018 (.015)	0.229	-.01, .05
Grade endometriosis						
2 (grade 2)	.845 (.916)	0.359	-.97, 2.66	-.512 (.850)	0.548	-2.20, 1.17
3 (grade 3)	.504 (.885)	0.570	-1.25, 2.26	-.506 (.821)	0.539	-2.13, 1.12
4 (grade 4)	.728 (.830)	0.383	-.92, 2.38	-.682 (.770)	0.378	-2.21, .85
BMI	.011 (.023)	0.631	-.03, .06	.024 (.022)	0.261	-.02, .07
Age categories						
2 (30-39 years)	-.266 (.296)	0.371	-.85, .32	-.194 (.274)	0.480	-.74, .35
3 (40-49 years)	-1.025 (.346)*	0.004	-1.71, -.34	-.319 (.321)	0.322	-.96, .32
4 (50-59 years)	-2.984 (.876)**	0.001	-4.72, -1.24	-3.165 (.812)**	0.000	-4.78, -1.55
Presence of infertility	.354 (.274)	0.199	-.19, .90	-.104 (.254)	0.684	-.61, .40
Constant	7.122 (1.082)	0.000	4.97, 9.27	10.623 (1.003)	0.000	8.63, 12.61
R ²	.220	-	-	.244	-	-
Adjusted R ²	.139	-	-	.165	-	-
Overall F-test	F = 2.71 (p=0.006)	-	-	F = 3.10 (p=0.002)	-	-

Note: * $p < 0.05$, ** $p < 0.01$

First, an ANOVA test was executed for the different grades of endometriosis. No substantial differences were examined in testing the median in a boxplot for these different grades of endometriosis. Thus, no differences were expected in the mean. The one-way ANOVA tests were statistically insignificant for the grade of endometriosis variable ($p=0.469$ and $p=0.654$, both > 0.05).

Another ANOVA test was executed for different age groups facing endometriosis in our sample. Medians seem to be substantially higher for age group 1 (18-29), 2 (30-39) and 3 (40-49) compared to the fourth age group (50-59) for both outcome variables. In the one-way ANOVA tests, statistically significant differences were elicited between the means for log yearly health care costs/log yearly total costs of the different age categories (p -value < 0.05). Due to the results of the Barlett's test for yearly health care costs (p -value= $0.397 > 0.05$), the null-hypothesis that variances are equal cannot be rejected. Thus, statistically significant differences in average annual health care costs seem to exist between different age categories. The Barlett's test for the annual total costs ($p=0.014 < 0.05$) indicated that we have to reject H_0 that variances are equal which means we cannot assume equal variance for annual total costs. Results for both ANOVA tests are stated in Table 11.

Table 11. Results ANOVA test for grade of endometriosis and age categories (n=107 women)

	Yearly health care costs (log)	Yearly total costs (log)
Grade endometriosis		
One-way ANOVA	p=0.469	p=0.654
Barlett's test	p=0.925	p=0.627
Age categories		
One-way ANOVA	p=0.006**	p=0.008**
Barlett's test	p=0.397	p=0.014*

Note: * $p < 0.05$, ** $p < 0.01$

Results focus group sessions

The distribution of the focus group sample may be skewed due to the fact that all participants had a long history with the illness. Also, long struggles with the diagnosis and treatments were present which may have led to a skewed sample population. Women who did participate possibly had a stronger urge to raise consciousness regarding the illness than women that did not participate. This also had a possible effect on the results. Women spoke openly and extensively and were motivated to tell about their prevailing experiences, pain and frustrations. One main underlying theme (i.e. “the long struggle”) was identified during the focus group sessions, consisting of seven subthemes: “uncertainty and day-to-day life”, “impact on daily life”, “frustration”, “awareness” and “maternity and infertility”. The extensive list of subthemes and categories associated with the main theme can be found in Table 12. Some categories will be discussed below and the most recurrent categories for all subthemes are stated in Table 12. Note that this is not an exhaustive list of categories examined during the focus group sessions, but the most recurrent categories examined by co-occurrence were included.

Table 12. Main theme, subthemes and categories of focus group analyses

Theme	Subtheme	Most recurrent categories
The long struggle	<i>Uncertainty and day-to-day life</i>	<ul style="list-style-type: none"> - Not knowing what is going on - Uncertainty on development of disease - Uncertainty in outcomes of surgeries and treatments - Uncertainty in maternity and infertility - Uncertainty in expression of illness; differs every day - Coping with endometriosis - Financial uncertainty (e.g. in terms of work) - Overall impact on daily life (i.e. emotional impact due to uncertainty)
	<i>Impact on daily life</i>	<ul style="list-style-type: none"> - Emotional/mental impact - Financial impact - Impact on lifestyle - Impact on relationships - Impact on social life - Impact on work and understanding (e.g. pushing limits) - Impact on study - Finding balance - Making choices to adapt daily life to endometriosis - Impact of medical treatments on daily life - Frustration - Uncertainty - Maternity and infertility - Deteriorated quality of life

<i>Frustration</i>	<ul style="list-style-type: none"> - Frustration associated with knowledge of relatives and physicians - Feeling misunderstood - Frustration associated with the experiences with treatments - Frustration associated with the struggle - Frustration associated with the impact on daily life
<i>Awareness, stigma and invisible illness</i>	<ul style="list-style-type: none"> - Little awareness on a societal level - Invisible illness - Stigma on endometriosis - Knowledge of relatives and physicians - Information on the disease - Ignorance of others

Note: This is not an exhaustive list of associated categories. However, main categories are presented here.

The long struggle

The long struggle was identified as the main underlying theme of the focus group sessions. This long struggle is defined as the challenges faced by women with endometriosis throughout the entire course of the disease. The struggle occurred before, during and after diagnosis. Most women experienced a long struggle of getting diagnosed after the first moment of having symptoms. This is in line with the baseline characteristics of the burden of illness study that showed an average of 12.2 years (SD: 7.43). Despite, most women indicated that all of their symptoms and complaints fell into place after they got diagnosed. Major components of this long struggle were severe pain and symptoms (e.g. excessive menstrual bleeding), visiting many doctors, frustration and uncertainty:

“I got my first period when I was 11. When I was 12, after I have been on my period for more than three months, I got “contractions”. At that moment, every physician understood something was wrong. Afterwards, I have been physically examined but nothing has been found. They sent me home and said: “You just have extreme periods. Don’t worry, when you’ve had your first child, this will all be over.” I was 12, so that wasn’t a very exhilarating message [...] When I was 38, I suffered from thrombosis and a double pulmonary embolism [...] When I visited the University hospital, during the first ultrasound scan, they noticed that endometrial tissue was everywhere.” (participant 5)

The emotional impact before getting the diagnosis and thus this long struggle were elicited to be very extensive. It made women in the focus group become harsh and even let them doubt about their own feelings and thoughts:

“You even start doubting yourself: am I the one that’s going crazy? Is everything just happening in my own head? Am I a whiner? Every single scenario crosses your mind, even though you know yourself that something is absolutely wrong.” (participant 7)

Overall, the long struggle of getting diagnosed seemed to be present for most participants. However, two exceptions were present in which one woman got diagnosed soon after getting symptoms and one by accident during her pregnancy. The latter also recognized all her symptoms of those preceding years after diagnosis:

“The endometrial tissue in my body was discovered somewhat by accident. At that moment, I was pregnant for 20 weeks. At once, I got severe abdominal pains. Two arteries of my uterus turned out to be destroyed by the endometrial tissue. [...] My symptoms already started when I was 12. (participant 2)

Also, extensive experiences with medical, lifestyle and alternative treatments during the process of diagnosis and afterwards were obviously present. Some women had already undergone many treatments while others just got diagnosed and were left in much uncertainty of what would happen in the future.

Not only the long struggle of getting diagnosed, but additionally the long struggle of uncertainty, the impact of endometriosis on daily life, frustration, lack of awareness and maternity and infertility were elicited during the sessions and are discussed below as subthemes. All subthemes identified here interact with each other and have some overlap. Pain seemed to be a major aspect that has been elicited in all aspects thoroughly and throughout the entire sessions. Collectively, they lead to the extensive burden of endometriosis.

Uncertainty and day-to-day life

Uncertainty seemed to be one of the key elements in daily life for women with endometriosis. This uncertainty was represented in not knowing what is going on in their bodies, uncertainty about the development of the disease, on outcomes of treatments and surgeries, not knowing what will happen in the future in terms of maternity and infertility and uncertainty around the expression of the illness that differs every day. Living a life on a day-to-day basis thus seemed to be needed in order to cope with endometriosis, because they could not make plans for the near or late future:

“I am dealing with endometriosis every second of the day. I’m constantly worrying about the surgery and its outcomes and if I will still be able to become pregnant. It’s really a rough time.” (participant 2)

“[...] balance is very hard to find. This is the case since every day is a different day. If the situation was the same every day, you could have learned to deal and to cope with it. But it’s not the same every day.” (participant 2)

Impact on daily life

Endometriosis seemed to be present every day as a struggle in the lives of the participants. This daily life impact was elicited in, for example, emotional impact, financial impact, impact on lifestyle, relationships, social life, study and work. Finding a balance in life and limiting these negative effects of the illness seemed easier said than done. Two women mentioned the complexity of the illness, how hard it is to deal with endometriosis and the accompanied powerlessness:

“It is all about looking for the right balance. How am I managing nutritional intake; what do I want to do hormone-wise (hormonal treatment); how can I make my life worthwhile living again; and how can I make myself happy again? Affecting social aspects, but also work-related. I like my work and I want to continue doing it as long as I can. But processing not being able to fulfill the desire to get children is also an important issue.” (participant 6)

One main category of impact on daily life elicited during the sessions was the emotional and mental impact. Two women directly mentioned suffering from a depression while the other five also indicated indirectly to suffer from the emotional impact. This was attributed to severe pain and was even to the extent that there was no joy of being alive anymore:

“I told them (the gynecologists): I can’t take it (the pain) anymore. I am currently considering a three-story drop (committing suicide), because I can’t take the pain anymore.” (participant 1)

This emotional impact was also due to experiences with treatments like contraception and the IUD that suppressed the emotions and feelings of three participants:

“I have been saving myself (from endometriosis) for 20 years by using contraceptives continuously. Now, I found out I have been acting substantially different than I actually am due to that (contraception). Darker, more negative thoughts, a decreased libido.” (participant 4)

For impact on lifestyle, the majority of the women indicated to be very tired and daily living had to be adapted to cope with this lack of energy. Almost all women indicated not having energy anymore to fulfill their social contacts. Additionally, the current COVID-19 measures seemed to give them some sort of relief since their regular social activities were now eliminated from their daily lives. This seemed to lead to some sort of relaxation. Only one of them indicated that she did not suffer much from social pressure and contrarily got energy from these activities (if pain was acceptable). However, many women seemed to push their limits in order to remain functioning well in daily life. This theme returned mainly in work:

“I have done this for many years: I kept on working continuously; loyalty towards my work and employer; setting standards higher and higher. Afterwards, I regretted pushing my limits extensively because I didn’t listen to my own body. At a certain moment, this caused my body to be fully exhausted due to keep on going and never stopping.” (participant 5)

Understanding at work seemed to be present to some extent. However, one woman indicated that there may be understanding, although there is no solution to the problem. This is a constant struggle and also coincides with pushing one’s limits. One woman stated that understanding at work consists of two parts: your employer and your co-workers. Co-worker understanding is related to reintegration, while understanding of the employer relates to the financial impact of the illness. Thus, financial impact seemed to be important as well, where a lack of financial support was elicited as a central topic:

“Now, I am considering several options: shall I follow a diet; I actually want to go to an osteopath; I am considering supplements. However, everything requires a lot of money. And if you don’t have that, it makes it very hard to choose between alternatives: am I going to buy a new pair of shoes or am I going to visit the osteopath twice for the same amount of money? Those choices are mad choices to make.” (participant 4)

Additionally, endometriosis seemed to affect relationships with their significant others. Most partners were indicated to be supportive. However, relationships with partners were associated with guilt and some women faced difficulties talking about this. Sexual relationships, processing together and feeling understood were main themes elicited in the sessions:

“[...] I said this (he could leave me if he wanted to) to my boyfriend many times. He has to accept a lot. He has to take a lot for granted: socially, energy-wise, sexually. Literally every aspect of a relationship is affected by endometriosis.” (participant 5)

Frustration

Another main identified subtheme was frustration. This frustration was namely present towards physicians that did not provide the right information or did not have proper knowledge on how to diagnose or treat women in the right way. Two women indicated having had good experiences with physicians, while the others did not. Mostly, women did not feel understood or heard and felt like whiners many times:

“That is mind-blowing. If a GP or an E.R. doctor doesn’t have proper knowledge on the illness, I may understand. However, it is inconceivable that many gynecologists still don’t have any knowledge on endometriosis. That’s just wrong. When this is your profession, you should know it (having knowledge on the disease).” (participant 3)

Not only physicians were a source of feeling misunderstood, also relatives, close friends, people at work and society as a whole:

“At work people always tell me: it could always be worse. At least you cannot die from it (endometriosis).” (participant 7)

“After diagnosis, people surrounding me immediately told me: “if you got your surgery, everything will be solved”. But no, that’s not the case. This (endometriosis) will be there for the rest of your life.” (participant 1)

Besides frustration induced by physicians, experiences with prescribed treatments were elicited to be a source of frustration as well. This frustration mostly arose towards the fact that standard treatments like hormonal therapy and diets (e.g. FODMAP-diet) were the first things women continuously got prescribed by physicians as being the solution since they reduce growth of endometrial tissue and may reduce abdominal-related symptoms respectively. However, these treatments are to manage pain and to control symptoms as no possible solution to the disorder exists [51]. Particularly, insufficient attention is given to what treatment suits best for the individual woman (e.g. preferences regarding the desire to have children and their body’s and mind’s reaction to certain therapies) which in turn caused frustration:

“That is the reason why I am disappointed. They (physicians) are focusing too much on that: “Start using contraceptives, start trying the FODMAP-diet”. However, it is just a plaster and nothing more than that.” (participant 1)

Awareness

Another main subtheme elicited in the focus groups was awareness of endometriosis. For endometriosis little awareness and information seemed to be present at society level:

“Endometriosis is even stated in the top 20 most painful illnesses. And still, so little is known.” (participant 1)

Accordingly, endometriosis was indicated as an invisible illness where women seem to have to deal with it mostly individually. From the outside, nobody can see or judge whether one is suffering from a chronic disease:

“A lot of people just don’t know. If you just see me now, there is no label on me. Nobody knows what is really happening to me. You do ask for attention, but only just a little. However, you also have a high tolerance for pain, so you just accept it, over and over again.” (participant 4)

Additionally, participants notified that a stigma is present on associated themes of endometriosis: a stigma on women just not being able to accept pain and on endometriosis not being a chronic disease but just “small period-related pains”. However, this is a serious condition for these women, which affects not only the inside of their uteruses, but also their bladders, colons, ovaries and fallopian tubes. These effects and pains seem to be assumed to be part of the deal for women. For example, many people, including women themselves, perceive pain as “being part of your period”, so one should not complain and just accept it:

“[...] pain, it is what it is: you are a woman, it’s part of the deal. That is exactly what they have been saying since we were little. We have become so hard for ourselves. We even lost the connection with our own bodies and forgot how to care for ourselves. All because you are constantly trying to survive, but at the same time you are not living at all.” (participant 1)

The combination of these three aspects, awareness, having an invisible illness and the existing stigma, seemed to lead to feeling misunderstood and not feeling heard.

Maternity and infertility

Maternity and infertility seemed to be one of the main aspects of the burden of endometriosis according to the focus group participants. As mentioned previously, a majority of the women in the focus groups got a late diagnosis which may have led to infertility. Even if they were able to get children, the process seemed to be a long struggle due to severe infertility trajectories, uncertainty and the extensive emotional impact. A major category in maternity and infertility is the emotional impact of not being able to get children:

“A good example is Mother’s Day, last Sunday. For me, it will always be a horrible day. It will always remain a running score.” (participant 5)

“To be honest, even though my desire to have children is tremendous, I am thinking about giving up that desire and having my uterus removed. Everything to not feel the pain of the adenomyosis.” (participant 1)

Four women literally indicated having lost an unborn child due to endometriosis, one with the risk of not surviving the critical situation herself. This whole process was not only affecting these women emotionally, but also processing with their partners was elicited to be tough. However, two women did get children at younger ages. Lasting gratitude for getting children was present in their stories and experiences. However, maternity and infertility seemed to be coinciding with the main theme of a long struggle, also for these women. Trade-offs were made between maternity and for example study, work and career perspective:

“The gynecologist told me: if you want to get pregnant you will have to take advantage of the years in which you are still fertile. So I became a mother at a young age, even though I possibly wanted to do other things for three more years, like work. Everything is affected by endometriosis.” (participant 4)

One woman indicated the enormous emotional impact of taking contraception every day (and sometimes even two different types of contraception simultaneously) as a treatment to suppress the growth of endometrial tissue, whilst also having the desire to get children. This was confirmed by others. At the same time, it felt very contradictory for these women:

“[...] every day it was like a hard pill to swallow. Every day, I took contraception to avoid having something that I desperately wanted to have, but what I was ultimately not able to have (a child). That was extremely hard to hear. I wanted to take every pill instead, but I already had psychological resistance towards contraception.” (participant 3)

One woman was diagnosed shortly after noticing the symptoms thus not experiencing a long diagnostic delay. However, she experienced a struggle in the process of getting children due to uncertainty, causing an enormous emotional impact:

“[...] I do have children, but the long road towards getting children has always been affected by endometriosis. I didn't know if I was able to get children. When I was 18, I had to tell my current husband “it may be possible that you have a girlfriend that is not able to get kids”.” (participant 4)

On the other hand, two women indicated they desperately wanted to have children and could not, but were secretly happy that they could not since they were not physically able to care for these children due to their condition:

“[...] sometimes I ask myself: what if I could get children? Where on earth does my energy have to come from to be able to care for this child as well?” (participant 6)

Discussion

This is the first mixed methods study in the Netherlands that investigated the burden of disease and experiences with endometriosis in daily life. This was done by means of a mixed methods explanatory sequential design. Here, survey data collection was followed by deriving descriptive statistics, costs and quality of life, executing two multiple regression analyses and determining the main cost drivers. Two focus group sessions examined the experiences, feelings and thoughts accompanying daily lives of women with endometriosis which complemented the results of the burden of illness study. This was analyzed by means of a deductive analysis.

Costs

Results for the cost determination show that annual total costs of €28,772 (95% CI: €23,276-€34,269) were present for women suffering from endometriosis, where €5,323 (95% CI: €4,027-€6,619) was represented by direct health care costs. These total costs were substantially higher than the costs reported in Simoens et al. (2012) [15] which amounted to €9,579 (95% CI: €8,560-€10,599) with only €3,113 (95% CI: €2,251-€3,975) represented by direct health care costs. In that study, not only distinct differences were examined in the results of costs, but also in the determination of the main cost drivers. Be aware of the fact that costs of Simoens et al. (2012) [15] were in 2010 Euros. If Dutch CPI-mutations were used, prices should be multiplied by 1.17 to get Dutch 2020 prices [41]. Applying these price mutations also

resulted in substantially lower costs in Simoens et al. (2012) [15] compared to our study. Another burden of illness study conducted in Europe (i.e. Belgium) was Klein et al. (2014) [6]. They calculated similar total costs and health care costs as Simoens et al. (2012) [15]. Compared to other major chronic diseases in the Netherlands, such as diabetes mellitus, costs are substantially lower than the outcomes of the burden of endometriosis in this study [52]. The estimated yearly (2016) direct health care costs of diabetes mellitus amounted to €1,429 and productivity costs amounted only to €590 per patient in the Netherlands [53].

The majority of the higher total costs in our study originated in the productivity loss reported by our study population. In this study, annual productivity costs represented approximately 70.6% of annual total costs, which amounted to €20,303 (95% CI: €15,389-€25,217). Simoens et al. (2012) [15] elicited productivity costs of €6,298 (95% CI: €5,826-€6,771; i.e. 65.7% of annual total costs), which are substantially lower. This large difference may be attributed to the friction costs method used in our study where short paid work absenteeism, paid work presenteeism and unpaid work presenteeism were multiplied by 13 to get annual productivity costs [40]. This is an assumption that may have caused an overestimation of the costs in our study. Remaining days within the friction period were accounted for in the analysis for women that were absent for longer than four weeks but shorter than the friction period. Zero-costs were attached to participants that were absent for longer than the friction period since vacancies were filled after this period and no further costs were present for society. Simultaneously, women with paid work were absent for on average 40.28 (SD: 177.77) days. This average may be high due to the 17 participants that were absent for longer than four weeks. These are outliers and drive up the average substantially. However, this may show the severeness of the absence of work in days. These women were not able to work which, from an individual perspective, is an adverse event in itself. This individual perspective is represented in the human capital method [54]. However, the friction period is used in this study according to the Dutch manual for costing studies and the chosen societal perspective [40].

Several factors could have influenced the results of this study and may explain the deviations from the existing literature base. Results for the baseline characteristics show that approximately 76% of the sample population, representing 81 women, indicated to suffer from comorbidities. Our survey consisted of medical care use, productivity loss and quality of life questions related to endometriosis. It could be possible that this has been misread and health care, costs and productivity loss related to other illnesses and comorbidities have entered the study. This may have caused an overestimation of the costs. However, other studies [55-57] also indicated the clear presence of comorbidities which thus may be part of the burden of endometriosis. This phenomenon should be considered in future research.

Additionally, the higher costs elicited in the burden of illness study may have been the result of the sampling technique via the social media posts of the Endometriosis Foundation of the Netherlands. It may be the case that women who suffer more severely are more affiliated with the foundation. They are more willing to seek information and updates and may be more motivated to raise awareness which possibly resulted in volunteer bias and sampling bias. This may have led to the skewed population sample towards grade four of endometriosis, representing women that are more sick. However, many women with grade one and grade two are never diagnosed with endometriosis which makes it hard to include this part of the population in the sample with the existing inclusion criteria of a clinical diagnosis [37]. The

results can thus show how bad the illness can be, but are not the representative averages of all women with endometriosis in the Netherlands. The Dutch figures show that the majority of women suffers from grade one or grade two (i.e. 90%) and only 10% from grade three or four of endometriosis [37]. This may have been a threat to the population aspect of the external validity since the sample of this study may not be representative for the true population. This may have led to sampling bias because the sample was not completely random. Collecting a more representative sample is recommended for future research in order to enhance the external validity of the study.

Comparing our sample to the sample of Simoens et al. (2012) [15] elicited different sample sizes, occupation characteristics and grades of endometriosis. Our sample size is relatively small, 107, compared to 909. Approximately 16% is unable to work in our sample compared to 9% in theirs and grade two and three represented 89% in our sample compared to 72% in theirs. However, these differences in percentages may just have been the result of our substantially smaller sample. This skewed population and outliers also resulted in non-normality in the two outcome variables. The log-transformation of the two outcome variables for the regression analysis accounted for this phenomenon.

Quality of life

Results for quality of life show that the EQ-5D-5L and MHQoL index values also contain negative minimum values, although zero theoretically represents death. This lower boundary may have been crossed since there is a possibility of valuing some situations as worse than death. The upper boundary was not and cannot be crossed [33]. The average EQ-5D-5L index value of 0.585 (SD: 0.277) seems to be substantially low compared to the general population index value in the Netherlands where the average for women is 0.858 (SD: 0.168) [33]. An average or overall endometriosis EQ-5D-5L index value is not examined yet in the Netherlands. Spronk et al. (2021) [58] have calculated the EQ-5D-5L index value for fatigue measures with and without chronic disease. The average fatigue EQ-5D-5L index value with a chronic disease amounts to 0.72 (SD) on average [58]. This is substantially higher than the elicited EQ-5D-5L index value in our sample (i.e. 0.585).

Average results for EHP-30 scores ranged between 53.80 (median: 53.33, SD: 15.18) for emotional wellbeing and 65.20 (median: 70, SD: 20.42) for control and powerlessness. Another Dutch study of van de Burgt, Hendriks and Kluivers (2010) [34] showed median EHP-30 scores between 33 for self-image and 54 for control and powerlessness. These values are lower than the results in this study. This indicated worse health-related quality of life for patients with endometriosis in our sample as compared to theirs. The average MHQoL index value amounted to 0.568 (SD: 0.318). This average cannot be compared to the overall population of the Netherlands since these utility tariffs are not available yet (i.e. the study presenting the utility tariffs has not been published). The elicited average MHQoL index score of 10.24 (SD: 2.14) is approximately equal to and even slightly lower than the average preliminary MHQoL index score of the mental health population of the Netherlands which on average amounts to 11.50 (SD: 4.00) [35].

Lastly, EQ-5D-5L and MHQoL index values seem to be correlated somewhat since they have some overlap in questions in for example daily activities and the mental health-related aspect. Higher EQ-5D-5L index values correlate with higher MHQoL index values and vice

versa. Direction of the correlation cannot be examined since there may be reverse causality. This correlation was further investigated by means of a sensitivity analysis in which the MHQoL index values were removed from the regression. The results confirmed an existing issue with these two variables. Future research should further investigate the implication of including both regressors, the EQ-5D-5L index value and the MHQoL index value, in one analysis.

Multiple regression analyses

For the inferential statistics, diagnostic delay, and age categories 40-49 and 50-59 were found to be significant determinants of the annual yearly health care costs at a 5% significance level. Age category 50-59 was the only determinant found to be significant for annual total costs at a 5% significance level. R^2 and adjusted- R^2 were substantially low. Although the F-statistics of the models were significant at a 5% significance level, many individual regressors were insignificant possibly due to the multicollinearity problem mentioned in the assumption testing. This makes it hard to separate out single effects [59,60]. In Simoens et al. (2012) [15], all corresponding independent variables showed a significant effect for annual health care costs. For annual total costs, the only corresponding independent variables that were significant in their study were age, the presence of infertility and annual number of QALYs. For QALYs, in addition to the utility values that represent quality of life, the duration of time in certain health states is incorporated [61]. Health-related quality of life and mental health-related quality of life included in our regression analyses are stated as health-utility values at a certain point in time [62].

The differences in results for the inferential statistics between the two studies could have been caused by several reasons. Firstly, these differences may be present due to different compositions of age groups in the sample or having included different variables in the regression analyses. Secondly, sample sizes were substantially different comparing the two studies, namely 909 as compared to 107 in this study. Thirdly, a backward stepwise entry mechanism for selecting the independent variables in their study compared to forced entry in this study may also explain the different results. Stepwise inclusion methods are based on small differences in semi-partial correlation which may be in contrast with the theoretical importance. Lastly, there were differences present across our sample populations, with the population of Simoens et al. (2012) [15] having less severe grades of endometriosis compared to our population. This may have led to the unexpected insignificant coefficients for other included independent variables in our study. On the contrary, Simoens et al. (2012) [15] collected respondents on being treated at least once in an endometriosis-specific center. This also may have led to a different composition of the sample. R^2 and adjusted- R^2 were of a somewhat similar size for the two studies which means the predictive power of the model is approximately the same.

Other reasons for not finding significant results may exist. For example, for infertility, only a small amount of women (i.e. 27 out of 107) indicated to be infertile which may have led to non-significance of that variable on the two cost-outcomes. This may be explained by the fact that in the data only five women who indicated to suffer from infertility also indicated to visit fertility specialists or undergo in vitro fertilization (IVF) or other fertility treatments. The other 22 women who indicated to be infertile did not report any visit or treatment. Additionally, five

women that did follow those trajectories did not indicate to suffer from infertility. This may have led to self-reporting bias. The author could not find any literature concerning this interesting result. Future research may further investigate this phenomenon.

Focus group sessions

For the focus group sessions, six main themes were elicited for the burden of endometriosis. The overarching theme was the long struggle from symptoms to diagnosis, but also after diagnosis to find balance in life with endometriosis. In turn, the long struggle consisted of uncertainty and day-to-day life and impact on daily life concerning negative consequences on different aspects (e.g. work, pain, emotionally, relationship with partner, social life). Frustration towards physicians and close friends/family (due to a lack of understanding and knowledge) and not feeling heard were prevailing as well. Endometriosis also seemed an invisible illness where a lack of awareness among society but also among women themselves exists. Lastly, maternity and infertility seemed to be a prevailing struggle in which women emotionally fight against the possible prognosis of not being able to get children. When comparing the qualitative results to the quantitative outcomes, similarities were found. The major and significant results from the burden of illness study for diagnostic delay (e.g. the long struggle), health-related quality of life (e.g. pain), mental health-related quality of life (e.g. emotional impact, depression), high health care costs (e.g. long struggle, many doctors), high productivity costs (e.g. work) and annual total costs seemed to be elicited in the focus group sessions as well. The high average score for the domain control and powerlessness in the EHP-30 was confirmed by the long struggle and uncertainty elicited in the focus group sessions. Also new concepts arose during qualitative research that were not examined in the quantitative study. This was caused by the more explorative nature of the focus groups and the fact that women could indicate themselves what specific effects endometriosis may have on their lives. An example of a new concept that arose from the focus groups was the effect of endometriosis on relationships.

The main results of the focus groups are similar to other studies: a long struggle towards diagnosis and after diagnosis were prevailing. This long struggle was accompanied by frustration, day-to-day life with endometriosis and finding balance in life like in Cox et al. (2003) [21] and Ballweg (1997) [19]. The stigma on normalizing the extreme pain during periods by others, the feeling of not being heard or being misunderstood and the emotional damage are prevailing substantially in Cox et al (2013) [21] and our study. Just like the missing empathy by some GP's, which was also present in the study of Cox et al. (2013) [21], but not in the study of Montague and Wood (1997) [63]. Impact on women's education (e.g. unable to attend school due to symptoms or being less productive) seemed to be less prevailing in our sessions compared to Moradi et al (2014) [17]. This may have been the case since 37% of the sample of Moradi et al. (2014) [16] was between 16-24 years compared to only 14% being aged between 18-29 years in our sample. The schooling age was thus more represented in their sample and education was reported to be the most prevailing problem for age group 16-24 in their results. Women in our sample may not have perceived the effect on education as important anymore because of their higher average age. This may have been caused by the recall period. Consequently, other above-mentioned themes became more prevailing in our study.

For the focus group sessions, saturation for the total population was not reached due to homogeneity in disease severity of the study sample. The majority of women in the Netherlands

suffers from grade one or grade two (i.e. approximately 90%) while only 10% suffers from grade three or four of endometriosis [37]. This may have led to different views and experiences reported by the participants compared to the overall endometriosis population in the Netherlands. As mentioned earlier, this non-saturation was present because the obtained sample from the survey was already skewed towards severe grades of endometriosis and due to time constraints. These women may suffer more severely, seriously want to increase awareness for endometriosis and may be more motivated to participate in research for this reason, which possibly resulted in volunteer bias. This may have led to more extreme experiences in the results and possibly to results for the burden of endometriosis that may not fully represent the true burden since these women may be part of a minority of the overall population. However, the results show how severe the illness can be. Future research may focus on a different sampling technique and recruitment procedure for the qualitative research, for example stratified sampling, to avoid such sampling problems in the future [64].

These severe outcomes may be very useful for enabling stakeholders to become involved in the process of suffering from endometriosis, for revealing underlying feelings and mechanisms and for uncovering the feelings of misunderstanding by creating awareness amongst stakeholders for the high burden of illness. Examples of stakeholders are hospitals, physicians, nurses, the government and its advisory committees, health insurers and friends and family [65,66].

Strengths and limitations

Several assumptions were made in designing this research paper which may have led to some uncertainty in the results and may have caused an over- or underestimation of the costs and the burden in this study. A majority of the assumptions relate to the costing method of the quantitative analysis. An important assumption is the use of reference prices. These prices are feasible when reliable sources are used, but represent averages which may be an over- or underestimation in specific cases [67]. For outpatient treatments, the average general 2014 price of €276 was used (and discounted) while after randomly checking the data, a major occurring treatment, IVF trajectory, already costs €1,200 per treatment according to the price guidelines of the HMC [68]. Due to lack of specific price information on all indicated outpatient treatments and due to a lack of sufficient data, reference prices were used. For medicine use, only hormonal treatment was included due to insufficient data on pain medication and other medication. Based on discussion with a gynecologist and endometriosis specialist, Noortje Paridaans, hormonal treatment costs were considered as the major part of the costs of endometriosis treatments. Even though data on other medicine use is missing in the analysis, additional costs would thus not have been extremely distinct. Also, cheapest prices for hormonal treatments were used according to the Dutch manual for costing studies which eventually may have underestimated the true costs [40]. Endometriosis specialists within the research team reported that many women try numerous medicines to find out what suits best for them. Noortje Paridaans, gynecologist and endometriosis specialist Endometriose in Balans, indicated that most women do not take the cheapest alternative of all available medicines. In general, using the conservative method for missing use, frequency or dose may have led to an underestimation of the costs. Incorporating all these aspects may have resulted in even higher annual health care and total

costs. However, the methods and assumptions in this study were implemented/executed according to the validated Dutch manual for costing studies [40].

Several limitations were present concerning the current COVID-19 circumstances. The survey was online in February and March 2021 in which a recall period of at maximum since August-September 2020 was included. During this period, a partial lockdown was present in the Netherlands [69]. Non-critical planned regular health care was postponed or e-health methods were applied nationally during this period [36]. An underestimation of health care costs may be present if women did not indicate their cellphone appointments due to current COVID-19 measures as being physician visits. Additionally, many surgeries and non-urgent treatments and appointments were postponed due to the current COVID-19 situation [69]. Noortje Paridaans, gynecologist and endometriosis specialist *Endometriose in Balans*, indicates an approximate 25% of non-emergency care that has been switched from physical to eHealth appointments and a 60-70% reduction in surgeries for their specialized endometriosis center located in The Hague. This may also have led to an underestimation of costs. However, little is known at this point in time about the exact numbers of postponed surgeries and appointments in the Netherlands. Consequently, the exact impact of the COVID-19 pandemic on the costs associated with the burden of endometriosis is hard to determine. Lastly, travel costs were accounted for as if COVID-19 was not present, which means visits to hospitals and centers were accounted for as physical appointments for the calculation of travel costs. Other assumptions and possible consequences are explored in Appendix 1.

There were several limitations concerning the survey design. This study relied on self-reported endometriosis diagnosis. Therefore, we assumed respondents truthfully answering the questions. Consequently, biases and validity issues may have been present. Due to the retrospective character of this study, recall bias may have been present. This could have led to an under- or overestimation of costs since participants cannot recall every health care-related appointment. The iMCQ [27] is validated to assess the recall of medical services over the used six month period; hence, it is assumed that the impact of the recall bias is minimized [70]. Using an online questionnaire may have led to sampling bias since only women that have internet-based knowledge were integrated in the study. The length of the survey (i.e. 20-25 minutes) may have led to a lower completion rate and thus to a smaller sample size. However, a pilot study was done and no problems associated with the length of the survey were reported.

Some limitations have been present concerning the set-up of the focus groups. Researcher bias may have been present in terms of confirmation of the data. However, being constantly aware of this possible bias and constantly objectively criticizing and discussing all results within the research group may have limited this threat [48]. Another limitation to this study is that the results were not additionally validated by sending the results to participants and asking them whether or not the results were representative for their experiences. Due to time constraints this additional part of the analysis did not take place. In an ideal research situation this would have been executed and may be an improvement for further research. However, respondents had the freedom to mention any additional comments at the end of the focus groups and data was discussed thoroughly within the research group.

Lastly, results may not be easily extrapolated to different countries since only women in the Netherlands were included. Though, this is not a limitation of the study since we were interested in the burden of illness in the Netherlands.

Conclusion

This mixed methods study highlighted the severeness of the burden of endometriosis concerning substantial costs on a societal level, especially in terms of productivity loss. Also, the results showed low values for health-related quality of life as well as mental health-related quality of life. For the focus group sessions, severe and difficult experiences associated with the illness were elicited for our study sample. These experiences seemed to support the results of the examined costs and quality of life. Future research is needed to further investigate this high burden of illness by also including other societal costs like leisure time. The latter was partly represented in the focus group sessions but may also be a valuable extension for the burden of illness study. Additionally, future research may be conducted with the inclusion of other grades of endometriosis in a mixed methods study to achieve a more representative sample. One way to reach patients with other grades of endometriosis may be cooperating with health insurers to investigate the average health care utilization to obtain a more random and heterogeneous sample. This will be needed to raise awareness and in turn to decrease the diagnostic delay that may lead to a decrease in this substantial burden of endometriosis.

Acknowledgement

The author would like to thank the Endometriosis Foundation of the Netherlands for contributing to the data input. Additionally, she acknowledges Leonie Bremmers and dr. Leona Hakkaart-van Roijen of the Erasmus University Rotterdam and Noortje Paridaans of Endometriose in Balans located in The Hague, who has been supporting the study entirely.

References

1. Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, et al. ESHRE guideline for the diagnosis and treatment of endometriosis. *Hum Reprod*. 2005 Oct;20(10):2698-704. doi: 10.1093/humrep/dei135.
2. Armour M, Lawson K, Wood A, Smith CA, Abbott J. The cost of illness and economic burden of endometriosis and chronic pelvic pain in Australia: A national online survey. *PLoS One*. 2019 Oct;14(10). doi: 10.1371/journal.pone.0223316.
3. Simoens S, Hummelshoj L, D'Hooghe T. Endometriosis: cost estimates and methodological perspective. *Hum Reprod Update*. 2007 Jul-Aug;13(4):395-404. doi: 10.1093/humupd/dmm010.
4. Endometriose in Balans. Endometriose in balans: Expertisecentrum in het ziekenhuis [Internet]. Endometriose in Balans; 2019 [cited 2021 Jun]. Available from: <https://endometrioseinbalans.nl/>.
5. Ballard K, Lowton K, Wright J. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. *Fertil Steril*. 2006 Nov;86(5):1296-301. doi: 10.1016/j.fertnstert.2006.04.054.

6. Klein S, D'Hooghe T, Meuleman C, Dirksen C, Dunselman G, Simoens S. What is the societal burden of endometriosis-associated symptoms? a prospective Belgian study. *Reprod Biomed Online*. 2014 Jan;28(1):116-24. doi: 10.1016/j.rbmo.2013.09.020.
7. Broersen S. Diagnose endometriose meestal pas na jaren [Internet]. *Medisch Contact*; 2019 Oct 4 [cited 2021 Mar]. Available from: <https://www.medischcontact.nl/nieuws/laatste-nieuws/nieuwsartikel/diagnose-endometriose-meestal-pas-na-jaren.htm>.
8. Rogers PA, D'Hooghe TM, Fazleabas A, Gargett CE, Giudice LC, Montgomery GW, et al. Priorities for endometriosis research: recommendations from an international consensus workshop. *Reprod Sci*. 2009 Apr;16(4):335-46. doi: 10.1177/1933719108330568.
9. Taylor HS. Endometriosis: a complex systemic disease with multiple manifestations. *Fertil Steril*. 2019 Aug;112(2):235-6. doi: 10.1016/j.fertnstert.2019.06.006.
10. Drost RM, Paulus AT, Ruwaard D, Evers SM. Inter-sectoral costs and benefits of mental health prevention: towards a new classification scheme. *J Ment Health Policy Econ*. 2013 Dec;16(4):179-86.
11. Tariq S, Woodman J. Using mixed methods in health research. *JRSM Short Rep*. 2013 Jun;4(6). doi: 10.1177/2042533313479197.
12. Risberg T, Sørbye S, Norum J, Wist E. Diagnostic delay causes more psychological distress in female than in male cancer patients. *Hum Reprod Update*. 1996 Mar-Apr;19(6):625-39. doi: 10.1093/humupd/dmt027.
13. Wtop news. Delays in Diagnosing Endometriosis Can Lead to Irreversible Damage [Internet]. 2017 Jun. Available from: <https://wtop.com/center-for-innovative-gyn-care/2017/06/delays-diagnosing-endometriosis-can-lead-irreversible-damage/>.
14. Boseley S, Glenza J, Davidson H. Endometriosis: the hidden suffering of millions of women revealed. *The Guardian*. 2015 Sep 27. Available from: <https://www.theguardian.com/society/2015/sep/28/endometriosis-hidden-suffering-millions-women>.
15. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum Reprod*. 2012 May;27(5):1292-9. doi: 10.1093/humrep/des073.
16. Moradi M, Parker M, Sneddon A, Lopez V, Ellwood D. The Endometriosis Impact Questionnaire (EIQ): a tool to measure the long-term impact of endometriosis on different aspects of women's lives. *BMC Womens Health*. 2019 May 14;19(1):64. doi: 10.1186/s12905-019-0762-x.
17. Moradi M, Parker M, Sneddon A, Lopez V, Ellwood D. Impact of endometriosis on women's lives: a qualitative study. *BMC Womens Health*. 2014 Oct 4;14:123. doi: 10.1186/1472-6874-14-123.

18. Soliman AM, Surrey E, Bonafede M, Nelson JK, Castelli-Haley J. Real-World Evaluation of Direct and Indirect Economic Burden Among Endometriosis Patients in the United States. *Adv Ther.* 2018 Mar;35(3):408-23. doi: 10.1007/s12325-018-0667-3.
19. Ballweg ML. Blaming the victim. The psychologizing of endometriosis. *Obstet Gynecol Clin North Am.* 1997 Jun;24(2):441-53. doi: 10.1016/s0889-8545(05)70312-0.
20. Ballweg ML. Impact of endometriosis on women's health: comparative historical data show that the earlier the onset, the more severe the disease. *Best Pract Res Clin Obstet Gynaecol.* 2004 Apr;18(2):201-18. doi: 10.1016/j.bpobgyn.2004.01.003.
21. Cox H, Henderson L, Andersen N, Cagliarini G, Ski C. Focus group study of endometriosis: struggle, loss and the medical merry-go-round. *Int J Nurs Pract.* 2003 Feb;9(1):2-9. doi: 10.1046/j.1440-172x.2003.00396.x.
22. Culley L, Law C, Hudson N, Denny E, Mitchell H, Baumgarten M, et al. The social and psychological impact of endometriosis on women's lives: a critical narrative review. *Hum Reprod Update.* 2013 Nov-Dec;19(6):625-39. doi: 10.1093/humupd/dmt027.
23. Jones G, Jenkinson C, Kennedy S. The impact of endometriosis upon quality of life: a qualitative analysis. *J Psychosom Obstet Gynaecol.* 2004 Jun;25(2):123-33. doi: 10.1080/01674820400002279.
24. Qualtrics. Qualtrics software. Version January-June 2021 [software]. Provo, Utah, USA. 2020 [cited 2021 Jan]. Available from: <https://www.qualtrics.com/>.
25. StataCorp. Stata Statistical Software: Release 16 [software]. TX: StataCorp LLC; 2019 [cited 2021 Jan]. Available from: <https://www.stata.com/>.
26. Larg A, Moss JR. Cost-of-illness studies: a guide to critical evaluation. *Pharmacoeconomics.* 2011 Aug;29(8):653-71. doi: 10.2165/11588380-000000000-00000.
27. Bouwmans C, Hakkaart-van Roijen L, Koopmanschap M, Krol M, Severens H, Brouwer W. Handleiding iMTA Medical Cost Questionnaire (iMCQ). Rotterdam: iMTA, Erasmus Universiteit Rotterdam, 2013.
28. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L. The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses. *Value Health.* 2015 Sep;18(6):753-8. doi: 10.1016/j.jval.2015.05.009.
29. Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. *Qual Life Res.* 2008 Apr;17(3):463-73. doi: 10.1007/s11136-008-9318-5.
30. Jones G, Kennedy S, Barnard A, Wong J, Jenkinson C. Development of an endometriosis quality-of-life instrument: The Endometriosis Health Profile-30. *Obstet Gynecol.* 2001 Aug;98(2):258-64. doi: 10.1016/s0029-7844(01)01433-8.

31. Erasmus University Rotterdam: Erasmus School of Health Policy & Management. The Mental Health Quality of Life Questionnaire (MHQoL). 2018.
32. Bouwmans C, Hakkaart-van Roijen L, Koopmanschap M, Krol M, Severens H, Brouwer W. Handleiding iMTA Productivity Costs Questionnaire (iPCQ). Rotterdam: iMTA, Erasmus Universiteit Rotterdam, 2013.
33. Versteegh MM, Vermeulen M, Evers MAAS, de Wit GA, Prenger R, EA. S. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016 Jun;19(4):343-52. doi: 10.1016/j.jval.2016.01.003.
34. van de Burgt TJ, Hendriks JC, Kluivers KB. Quality of life in endometriosis: evaluation of the Dutch-version Endometriosis Health Profile-30 (EHP-30). Fertil Steril. 2011 Apr;95(5):1863-5. doi: 10.1016/j.fertnstert.2010.11.009.
35. van Krugten F, Busschbach J, Versteegh M, Hakkaart-van Roijen L, Brouwer W. The Mental Health Quality of Life Questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. Submitted.
36. Rijksoverheid. Extra maatregelen in ziekenhuizen door stijgende COVID-druk [Internet]. 2020 Dec [cited 2021 Jun]. Available from: <https://www.rijksoverheid.nl/ministeries/ministerie-van-volksgezondheid-welzijn-en-sport/nieuws/2020/12/21/extra-maatregelen-in-ziekenhuizen-door-stijgende-covid-druk>.
37. Endometriose in Balans. Diagnose [Internet]. 2019 [cited 2021 Jun]. Available from: <https://endometrioseinbalans.nl/diagnose/>.
38. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009 Nov;41(4):1149-60. doi: 10.3758/BRM.41.4.1149.
39. Dunet D. CDC Coffee Break: Introduction To Economic Evaluation 2012 Jan 10 [cited 2021 Jun]. Available from: https://www.cdc.gov/dhdsp/pubs/docs/cb_january_10_2012.pdf.
40. Hakkaart-van Roijen L, Van der Linden N, Bouwmans C, Kanters T, Tan S. Costing manual: Methodology of costing research and reference prices for economic evaluations in healthcare [in Dutch: Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg]. 2015.
41. CBS Statline. Consumentenprijzen; prijsindex 2015=100 [Internet]. 2021 [cited 2021]. Available from: <https://opendata.cbs.nl/#/CBS/nl/dataset/83131NED/table?ts=1623312609655>.
42. Zorginstituut Nederland. Medicijnkosten.nl [Internet]. 2021 [cited 2021]. Available from: <https://www.medicijnkosten.nl/>.
43. Zorginstituut Nederland. Farmacotherapeutisch Kompas [Internet]. 2021 [cited 2021]. Available from: <https://www.farmacotherapeutischkompas.nl/>.

44. UCLA: Institute for Digital Research & Education. Regression With Stata Chapter 2 – Regression Diagnostics [Internet]. 2021. Available from: <https://stats.idre.ucla.edu/stata/webbooks/reg/chapter2/stata-webbooksregressionwith-statachapter-2-regression-diagnostics/>.
45. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care*. 2007 Dec;19(6):349-57. doi: 10.1093/intqhc/mzm042.
46. Zoom Video Communications Inc. Zoom. Version 5 [software]. 2016 [cited 2021]. Available from: <https://zoom.us/>.
47. Peterson J. Presenting a Qualitative Study: A Reviewer’s Perspective. *Gift Child Q*. 2019 May;63(3). doi: 10.1177/0016986219844789.
48. Shah S. 7 Biases to avoid in qualitative research [Internet]. 2019 Jan 3 [cited 2021]. Available from: <https://www.editage.com/insights/7-biases-to-avoid-in-qualitative-research>.
49. ATLAS.ti Scientific Software Development GmbH. ATLAS.ti: qualitative data analysis. Version 9 Mac [software]. Berlin: ATLAS.ti Scientific Software Development GmbH; 1993. Available from: <https://atlasti.com/>.
50. Nuffic The Dutch Organisation for Internationalisation in Education. Education in the Netherlands [Internet]. [cited 2021 May]. Available from: <https://www.nuffic.nl/en/subjects/study-holland/education-netherlands>.
51. Endometriose in Balans. Behandelingen [Internet]. 2019 [cited 2021 Jun]. Available from: <https://endometrioseinbalans.nl/behandelingen/>.
52. Volksgezondheidszorg.info. Ranglijst aandoeningen op basis van vóórkomen [Internet]. 2021 [cited 2021 Jun]. Available from: <https://www.volksgezondheidszorg.info/ranglijst/ranglijst-aandoeningen-op-basis-van-vóórkomen>.
53. Peters ML, Huisman EL, Schoonen M, Wolffenbuttel BHR. The current total economic burden of diabetes mellitus in the Netherlands. *Neth J Med*. 2017 Sep;75(7):281-97.
54. Trovato M. Human capital approach in the economic assessment of interventions for the reduction of seismic vulnerability in historic centres. *Sustainability*. 2020 Sep;12(19). doi: 10.3390/su12198059.
55. Soliman AM, Yang H, Du E, Wu E, Castelli-Haley J, Winkel C. Comorbidities and symptoms among endometriosis patients: A systematic literature review. *Value Health*. 2016 May;19(3). doi: 10.1016/j.jval.2016.03.1445.
56. Teng SW, Horng HC, Ho CH, Yen MS, Chao HT, Wang PH, et al. Women with endometriosis have higher comorbidities: Analysis of domestic data in Taiwan. *J Chin Med Assoc*. 2016 Nov;79(11):577-82. doi: 10.1016/j.jcma.2016.04.006.

57. Soliman AM, Surrey E, Johnson S, Davis M, Castelli-Haley J, Snabes M. Incidence of comorbidities among women with endometriosis: a retrospective matched cohort study. *Fertil Steril*. 2016 Sep;106(3). doi: 10.1016/j.fertnstert.2016.07.796.
58. Spronk I, Polinder S, Bonsel GJ, Janssen MF, Haagsma JA. The relation between EQ-5D and fatigue in a Dutch general population sample: an explorative study. *Health Qual Life Outcomes*. 2021 Apr 29;19(1):135. doi: 10.1186/s12955-021-01771-3.
59. Field A. *Discovering Statistics Using SPSS*. Third ed. London: SAGE Publications Ltd; 2009. p. 197-260
60. StackExchange. Why is it possible to get significant F statistic ($p < .001$) but non-significant regressor t-tests? [Internet]. 2017 [cited 2021 Jun]. Available from: <https://stats.stackexchange.com/questions/3549/why-is-it-possible-to-get-significant-f-statistic-p-001-but-non-significant-r>.
61. Whitehead S, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull*. 2010 Dec;96(1):5-21. doi: doi.org/10.1093/bmb/ldq033.
62. Brown DS, Jia H, Zack MM, Thompson WW, Haddix AC, Kaplan RM. Using health-related quality of life and quality-adjusted life expectancy for effective public health surveillance and prevention. *Expert Rev Pharmacoecon Outcomes Res*. 2013 Aug;13(4):425-7. doi: 10.1586/14737167.2013.818816.
63. Montague M, Wood R. Diagnosing endometriosis: GP's knowledge, attitudes and views on diagnosis. Report to the Department of Health and Family Services General Practice Evaluation Program 1997 GPEP project. 1997:358.
64. Barratt H, Shantikumar S. Methods of sampling from a population [Internet]. *Health Knowledge*; 2009 [updated 2018; cited 2021 Jun]. Available from: <https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/methods-of-sampling-population>.
65. Kahan J. Focus Groups As a Tool for Policy Analysis. *Anal Soc Issues Public Policy*. 2001 Dec;1(1):129-46. doi: 10.1111/1530-2415.00007.
66. National Guideline Alliance (UK). Endometriosis: diagnosis and management. 2017 Sep; NICE Guideline, No. 73. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK453273/>.
67. Drost RM, Paulus AT, Ruwaard D, Evers SM. Valuing inter-sectoral costs and benefits of interventions in the healthcare sector: methods for obtaining unit prices. *Expert Rev Pharmacoecon Outcomes Res*. 2017;17 Feb(1):77-84. doi: 10.1586/14737167.2016.1141679.
68. Haaglanden Medisch Centrum. 2021 Passantentarieven Haaglanden Medisch Centrum: DBC Zorgproducten en OZP's. 2021.
69. Rijksoverheid. Situatie rondom corona blijft ernstig, lockdown wordt verlengd: Rijksoverheid; 2021 Feb 2 [cited 2021 Jun]. Available from: <https://www.rijksoverheid.nl/actueel/nieuws/2021/02/02/situatie-rondom-coronavirus-blijft-ernstig-lockdown-wordt-voortgezet>.

70. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc.* 2016;9:211-7. doi: 10.2147/JMDH.S104807.
71. Independer. [Internet]. [cited 2021 Mar]. Available from: www.independer.nl.
72. CBS Statline. Vacatures; SBI 2008; naar economische activiteit en bedrijfsgrootte [Internet]. 2021 [cited 2021]. Available from: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80472ned/table?ts=1624474239919>.
73. Frost J. Estimating a Good Sample Size for Your Study Using Power Analysis 2018 [cited 2021 Jun]. Available from: <https://statisticsbyjim.com/hypothesis-testing/sample-size-power-analysis/>.

Appendices

Appendix 1. Assumptions of the study

Table 13. Assumptions of the study

Subject	Assumptions	Possible consequences/threat
Specialist visits	Unrelated visits: some unrelated visits were identified and removed from the data after extensive discussion with gynecologist and endometriosis specialist Noortje Paridaans, Haaglanden Medical Center, The Hague (e.g. cardiologist, rheumatologist, oncologist, ENT specialist, hematologist and immunologist).	N/A
Double counting	Treatments-specialists: follow up treatments are removed for the sake of double counting since we assumed that they were already present in clinical specialist visits. This happened 6 times. In some cases it was unclear whether or not it would have been double counting. These treatments remained in the dataset.	Leads to uncertainty in the dataset.
Medication	Inclusion hormonal therapy: since many mistakes were made in answering the medicine use questions in the survey, only hormonal therapy is used. Those are the most substantive costs. This was discussed with endometriosis specialist Noortje Padriaans, Haaglanden Medical Center, The Hague.	There may be a small underestimation of medication costs. However, most substantive amount is accounted for in hormonal therapy, according to gynecologist and endometriosis specialist Noortje Paridaans, Haaglanden Medical Center, The Hague.
	Hormonal treatment prices: Hormonal treatment dosages and prices are used from medicijnkosten.nl [42] and Farmacotherapeutisch Kompas [43]. Cheapest prices for hormonal treatment were used according to the Dutch manual for costing studies [40].	This may lead to an underestimation of the true costs since many women do not use the cheapest alternative according to endometriosis specialist Noortje Padriaans, Haaglanden Medical Center, The Hague.
	Conservative methods medication use: missing use, frequency or dose in the data are incorporated as zero costs.	This may lead to an underestimation of the true costs.
	Hormonal therapy pharmaceutical costs: this medication is not OTC (Over The Counter). It has to be bought at the pharmacist with prescription of a doctor. Pharmacist costs are included. Since there is no data available on how many times patients get their prescription, a standard number for each hormonal therapy is used. Distinction is made between chronic use (oral contraceptive, oral progestins and GNRH analogues) and first delivery (IUD). This was discussed with Noortje Padriaans, Haaglanden Medical Center, The Hague.	Using a standardized assumed number for each hormonal therapy may underestimate pharmacist costs.
	Pharmaceutical care: guidelines and assumptions of the Dutch manual for costing studies were used to decide upon pharmaceutical costs and to make a distinction between first visit and chronic use [40].	It may under- or overestimate the true pharmaceutical costs.
Distance treatment centers	Hospital visits: ED visits, number of treatments hospital appointments with clinical specialist and overnight stays are included.	N/A
	Average price assumption: people mentioned way of travelling, but not how many times they used what way of travelling for what visit. Average price of ways of transport is thus used for people indicating more than one way of transporting.	It is an average and may not represent the complete true values. However, these differences may not be substantial.
	Single visits: assumed that people did not combine specialist visits or visits with treatments etcetera.	This may lead to a small overestimation since patients may combine visits and treatments.

	Physical therapist/GP: assumed people went with car/public transport. Most common way of travelling and recommended by the Dutch manual for costing studies [40].	It is an average and may not represent the complete true values. However, these differences may not be substantial.
	Distance and price for other facilities: average distance is assumed for all other facilities since this is the most logical distance out of the Dutch manual for costing studies [40].	It is an average and may not represent the complete true values. However, these differences may not be substantial.
Travel costs - general	Physical appointments: assumed that all appointments took place physically to account for travel costs (even though current measures may indicate otherwise).	May not represent the current situation of COVID-19 in which many visits were not physically (according to endometriosis specialist Noortje Padriaans, Haaglanden Medical Center, The Hague). However, it coincides with the regular patterns of travel costs.
Prices	Reference prices: reference prices of Dutch manual for costing studies were used for valuation of the costs [40].	It may under- or overestimate the true prices of cost determinants in the burden of illness study.
	CPI used: the most recent CPI is used [41]. Prices are thus for 2020 and not 2021 which eventually represents costs for 2020 and not 2021. Just like the used open and filled vacancies used in the friction cost method.	N/A
	Treatments: prices of all specific source of treatments are not available. Not reasonable to look at different hospitals for prices for every single treatment and take average of that. Thus we assumed the reference price of outpatient treatments mentioned in the Dutch manual for costing studies [40].	This may underestimate the costs. When looking randomly at average visits, most treatments take place as fertility treatment. Price of one treatment is extensively higher than the reference price used.
	Unavailable treatment prices in Dutch manual for costing studies: investigated different prices and average is taken. Dietician: all prices were approximately between €63-€73 per hour. Average is €68 [71]. Every appointment depends on time; mostly, there are visits of 15, 30, 45 and 60 minutes. Average time is assumed which leads to an average price of €34 per visit. Same process for homeopath. Occupational physician: comparable service is GP with a longer time span. Price of GP is multiplied by 2.	May not be an exact representation of the true costs. It is an approximation. This may under- or overestimate the true costs. However, this concerns just small differences.
	Psychologist: in absence of information on what kind of psychologist women attended, price of the first line psychologist is used.	This is the cheapest option, so this may underestimate the true costs for psychologist visits.
	Physical therapist, Cesar and Mensendieck remedial therapy: all have a price ranging from 30 to 34 euros. Physical therapist price of 33 euros is assumed.	N/A
Friction cost methos	Friction cost period: of 2020 is used. This friction period is calculated based on open and fulfilled vacancies in 2020 [72]. Further friction cost calculation are done by instruction of the Dutch manual for costing studies [40].	Information on yearly average vacancies is not yet available for 2021. This is the best estimation.
MHQoL	Preliminary tariff: is used for the MHQoL valuation. Obtained by personal communication.	N/A
	Outlier positive value one state: since it is possible to get a possible value for one state, it is possible to obtain a total utility value above 1. This is unreasonable and we assumed to be the MHQoL value to have a maximum of 1.	N/A

Appendix 2. Do-file data cleaning

```
** Conducted using Stata version 16.1

** drop not needed variables
drop StartDate
drop EndDate
drop ResponseType
drop Durationinseconds
drop Finished
drop DistributionChannel
drop UserLanguage
encode ResponseID, gen(ID)
drop ResponseID
order ID
label drop ID

** rename all variables

** general
ren DoelvanhetonderzoekHetdoel written_consent
ren Bentu18jaarofouder age18plus
ren Iserendometriosebijnvastges clinical_diagnosis
ren Welkegradatieinernstvandee grade_endo
ren Waarisuwendometrioseaangetro area_other_specification
ren Heeftunaastendometrioseeen comorb_otherspecification
ren Wanneerhaduvoorheteerstsym year_symptoms
ren Wanneerwerduwendometriosevoo year_diagnosis
ren AM diagnostic_delay
ren Totwelkeleeftijdsgroepbehoort age_cat
ren WatisuwgewichtInkilogramme weight
ren WatisuwlengteIncentimeters height
ren Watisuwburgerlijkestaat marital_status
ren Watisdehoogsteopleidingdie educ_level
ren AS educ_level_other
ren Watdoetuinhetdagelijksleve occupation
ren AU unable_work_perc
ren AV occupation_other

** iMCQ
ren Heeftuindeafgelopen6maande GP_GPnurse_yesno
ren Hoeveelafsprakenhaduindeaf GP_yes_POH_yes
ren AY GP_visits
ren AZ POH_visits
ren BA social_worker_visits_yesno
ren BB social_worker_visits
ren BC phys_therap_yesno
ren BD phys_therap_visits
ren BE dietician_yesno
ren BF dietician_visits
ren BG homeopath_yesno
ren BH homeopath_visits
ren BI psychologist_yesno
ren BJ psychologist_visits
ren BK occupational_phys_yesno
ren BL occupational_phys_visits
ren BM medicine_yesno
ren Welkemedicijnenheeftuindea paracetamol_per_use
ren BO paracetamol_use_per_day
ren BP paracetamol_days_per_6months
ren BQ NSAID_per_use
ren BR NSAID_use_per_day
ren BS NSAID_days_per_6months
ren BT Tramadol_per_use
ren BU Tramadol_use_per_day
ren BV Tramadol_days_per_6months
```

ren BW Oxynorm_per_use
ren BX Oxynorm_use_per_day
ren BY Oxynorm_days_per_6months
ren BZ Oxycontin_per_use
ren CA Oxycontin_use_per_day
ren CB Oxycontin_days_per_6months
ren CC Gabapentine_per_use
ren CD Gabapentine_use_per_day
ren CE Gabapentine_days_per_6months
ren HormonaletherapieOraleantic oral_contraceptive
ren HormonaletherapieHormoonhoud hormonal_IUD
ren HormonaletherapieOraleproge oral_progestativa
ren HormonaletherapieGNRHanaloo GNRH_analoog
ren HormonaletherapieOverignam hormonal_ther_other
ren Hoevaakbentuindefgelopen SEH_yesno
ren CU visits_SEH
ren CV ambulance_yesno
ren CW ambulance
ren Haduindeafgelopen6maanden hospital_appointment_yesno
ren Bentuindefgelopen6maanden outpatient_treatment_yesno
ren EA inpatient_treatment_yesno
ren Hoevaakheeftuindefgelopen hospital_overnight_times
ren Hoelangheeftuinhetziekenhu hospital_overnight_days
ren ED help_due_symptoms_yesno
ren Hoeveelwekenheeftudezehulp weeks_domestichelp
ren EG weeks_nursing
ren EH weeks_practicalhelp
ren Hoeveeluurhulpkreeguindeze hours_domestichelp
ren EJ hours_nursing
ren EK hours_practicalhelp
ren Welkewijzevanvervoerheeftu transport_other_specification
ren Watwasdeenkelereisafstandtu distance_hospital_km

** iPCQ

ren Hebtubetaaldwerk paid_work_yesno
ren HoeveeluurperweekwerktuTe paid_work_hours
ren Ophoeveeldagenindeweekwerk paid_work_days
ren Bentuindefgelopen4wekenabsenteism_yesno
ren EZ absenteism_days
ren Wasulangerdandegheleperio absenteism_long
ren Wanneerheeftuzichziekgemeld reportingsick_month
ren FC reportingsick_day
ren FD reportingsick_year
ren Warenerindeafgelopen4weken presenteeism_yesno
ren Ophoeveelwerkdagenhadutijde presenteeism_days
ren FH score_productiv_workwhensymptoms
ren Warenerdagenwaaropumindero missed_unpaid_work_yesno
ren OphoeveeldagenwasditzoTel missed_unpaid_work_days
ren Steldatiemandbijvoorbeelduw hours_replace_unpaidw_family

** EQ-5D-5L

ren MobiliteitZeteenkruisjeinhe EQ_mobility
ren ZelfzorgZeteenkruisjeinhet EQ_selfcare
ren Dagelijkseactiviteitenbijvwe EQ_daily_activ
ren PijnongemakZeteenkruisjein EQ_pain
ren AngstsomerheidZeteenkruisje EQ_fear
ren Wewillenwetenhoegoedofslec EQ_VAS_score

** EHP-30

ren Gedurendedelaatste4wekenho EHP30_skipsocials
ren FS EHP30_nohousehold_act
ren FT EHP30_pain_standing
ren FU EHP30_pain_sitting
ren FV EHP30_pain_walking
ren FW EHP30_no_sports
ren FX EHP30_no_appetite
ren FY EHP30_sleepless

```
ren FZ EHP30_bed_due_pain
ren GA EHP30_unable_dowhatwant
ren GB EHP30_unbearable_pain
ren GC EHP30_notfeelingwell
ren GD EHP30_frustrated_no_improvements
ren GE EHP30_frustrated_no_control
ren GF EHP30_unforgettable_symptoms
ren GG EHP30_symptoms_control_life
ren GH EHP30_symptoms_hinder_life
ren GI EHP30_depressed
ren GJ EHP30_sad
ren GK EHP30_miserable
ren GL EHP30_moodswings
ren GM EHP30_moody
ren GN EHP30_aggressive
ren GO EHP30_unable_tell_feelings
ren GP EHP30_felt_misunderstood
ren GQ EHP30_felt_whiner
ren GR EHP30_felt_lonely
ren GS EHP30_not_desired_clothes
ren GT EHP30_influenced_appearance
ren GU EHP30_no_confidence
```

```
** MHQoL
ren ZelfbeeldGeefaanwelkeuitspr MHQoL_selfimage
ren OnafhankelijkheidBijvoorbeeld MHQoL_independence
ren StemmingGeefaanwelkeuitspra MHQoL_mood
ren RelatiesBijvoorbeeldpartner MHQoL_relationships
ren DagelijkseactiviteitenBijvoorb MHQoL_dailyactiv
ren LichamelijkegezondheidGeefaa MHQoL_physhealth
ren ToekomstGeefaanwelkeuitspra MHQoL_future
ren HD MHQoL_psychological_wellbeing
```

```
** label all variables
```

```
** general
```

```
label var ID "respondent identifier"
```

```
label var written_consent "Geef u toestemming om uw antwoorden te gebruiken voor onderzoek?"
```

```
label var age18plus "Bent u 18 jaar of ouder?"
```

```
label var clinical_diagnosis "Is er endometriose bij u vastgesteld d.m.v. een kijkoperatie, MRI-scan of echografisch onderzoek?"
```

```
label var grade_endo "Welke gradatie in ernst van de endometriose is er d.m.v. deze onderzoeken bij u vastgesteld?"
```

```
label var area_peritoneum "Is er endometriose aangetroffen als spots op het buikvlies?"
```

```
label var area_bladder "Is er blaas endometriose aangetroffen?"
```

```
label var area_colon "Is er darm endometriose aangetroffen?"
```

```
label var area_ovaries "Is er endometriose in en rondom de eierstokken aangetroffen?"
```

```
label var area_adenomyose "Is er adenomyose aangetroffen (endometriose in de baarmoederwand)?"
```

```
label var area_vagina "Is er endometriose aangetroffen in/bij uw vagina?"
```

```
label var area_ureter "Is er endometriose aangetroffen in/bij uw urineleiders?"
```

```
label var area_unknown "Is de plaats waar uw endometriose is aangetroffen onbekend?"
```

```
label var area_other "Is uw endometriose op een andere plek aangetroffen?"
```

```
label var area_other_specification "Waar is uw endometriose aangetroffen? - Andere locatie dan bovenstaande opties, namelijk"
```

```
label var comorb_none "Heeft u naast endometriose geen enkele andere aandoening/ziekte?"
```

```
label var comorb_infertility "Bent u naast endometriose ook onvruchtbaar?"
```

```
label var comorb_migraine "Heeft u naast endometriose ook migraine?"
```

```
label var comorb_depression "Heeft u naast endometriose ook last van een depressie?"
```

```
label var comorb_backpain "heeft u naast endometriose ook rugklachten?"
```

```
label var comorb_fibroiduterus "Heeft u naast endometriose ook een vleesboom in de baarmoeder?"
```

```
label var comorb_asthma "Heeft u naast endometriose ook last van astma?"
```

```
label var comorb_thyroid "Heeft u naast endometriose ook een schildklier aandoening?"
```

```
label var comorb_fatigue "heeft u naast endometriose ook last van chronische vermoeidheid?"
```

```
label var comorb_PCOS "heeft u naast endometriose ook Polycysteus Ovarium Syndroom (PCOS)?"
```

```
label var comorb_other "Heeft u naast endometriose nog een andere aandoening dan boven genoemde?"
```

```
label var comorb_otherspecification "Heeft u, naast endometriose, een andere aandoening/ziekte? - Anders, namelijk"
```

```
label var year_symptoms "Wanneer had u voor het eerst symptomen van endometriose?"
```

label var year_diagnosis "Wanneer werd uw endometriose voor het eerst met zekerheid vastgesteld door een kijkoperatie, MRI-scan of echografisch onderzoek?"
label var diagnostic_delay "Wat is de vertraging in diagnose (diagnose-klachten)?"
label var age_cat "Tot welke leeftijdsgroep behoort u?"
label var weight "Wat is uw gewicht (kg)?"
label var height "Wat is uw lengte (cm)?"
label var marital_status "Wat is uw burgerlijke staat?"
label var educ_level "Wat is de hoogste opleiding die u heeft afgemaakt?"
label var educ_level_other "Wat is de hoogste opleiding die u heeft afgemaakt? - Ik heb een andere opleiding afgemaakt, namelijk"
label var occupation "Wat doet u in het dagelijks leven?"
label var occupation_other "Wat doet u in het dagelijks leven? - Ik doe iets anders, namelijk"
label var unable_work_perc "Ik ben arbeidsongeschikt voor ... %."

** iMCQ

label var GP_GPnurse_yesno "Heeft u in de afgelopen 6 maanden afspraken gehad met de huisarts of praktijkondersteuner (POH)?"
label var GP_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met uw huisarts?"
label var POH_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met uw praktijkondersteuner (POH)?"
label var social_worker_visits_yesno "Heeft u in de afgelopen 6 maanden afspraken gehad met een maatschappelijk werker?"
label var social_worker_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met een maatschappelijk werker?"
label var phys_therap_yesno "Heeft u in de afgelopen 6 maanden afspraken gehad met een fysiotherapeut? Of met een caesartherapeut, therapeut mensendieck of manueel therapeut?"
label var phys_therap_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met een fysiotherapeut? Of met een caesartherapeut, therapeut mensendieck of manueel therapeut?"
label var dietician_yesno "Heeft u in de afgelopen 6 maanden een afspraak gehad met een diëtist?"
label var dietician_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met een diëtist?"
label var homeopath_yesno "Heeft u in de afgelopen 6 maanden een afspraak gehad met een homeopaat of acupuncturist?"
label var homeopath_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met een homeopaat of acupuncturist?"
label var psychologist_yesno "Heeft u in de afgelopen 6 maanden een afspraak gehad met een psycholoog?"
label var psychologist_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met een psycholoog?"
label var occupational_phys_yesno "Heeft u in de afgelopen 6 maanden een afspraak gehad met de bedrijfsarts?"
label var occupational_phys_visits "Hoeveel afspraken had u in de afgelopen 6 maanden met de de bedrijfsarts?"
label var medicine_yesno "Heeft u in de afgelopen 6 maanden medicijnen gebruikt?"
label var paracetamol_per_use "Hoeveel paracetamol heeft u in de afgelopen 6 maanden per keer genomen?"
label var paracetamol_use_per_day "Hoe vaak op een dag heeft u in de afgelopen 6 maanden paracetamol genomen?"
label var paracetamol_days_per_6months "Op hoeveel dagen in de afgelopen 6 maanden heeft u paracetamol gebruikt?"
label var NSAID_per_use "Hoeveel NSAID's heeft u in de afgelopen 6 maanden per keer genomen?"
label var NSAID_use_per_day "Hoe vaak op een dag heeft u in de afgelopen 6 maanden NSAID's genomen?"
label var NSAID_days_per_6months "Op hoeveel dagen in de afgelopen 6 maanden heeft u NSAID's gebruikt?"
label var Tramadol_per_use "Hoeveel Tramadol heeft u in de afgelopen 6 maanden per keer genomen?"
label var Tramadol_use_per_day "Hoe vaak op een dag heeft u in de afgelopen 6 maanden Tramadol genomen?"
label var Tramadol_days_per_6months "Op hoeveel dagen in de afgelopen 6 maanden heeft u Tramadol gebruikt?"
label var Oxynorm_per_use "Hoeveel Oxynorm heeft u in de afgelopen 6 maanden per keer genomen?"
label var Oxynorm_use_per_day "Hoe vaak op een dag heeft u in de afgelopen 6 maanden Oxynorm genomen?"
label var Oxynorm_days_per_6months "Op hoeveel dagen in de afgelopen 6 maanden heeft u Oxynorm gebruikt?"
label var Oxycontin_per_use "Hoeveel Oxycontin heeft u in de afgelopen 6 maanden per keer genomen?"
label var Oxycontin_use_per_day "Hoe vaak op een dag heeft u in de afgelopen 6 maanden Oxycontin genomen?"
label var Oxycontin_days_per_6months "Op hoeveel dagen in de afgelopen 6 maanden heeft u Oxycontin gebruikt?"
label var Gabapentine_per_use "Hoeveel Gabapentine heeft u in de afgelopen 6 maanden per keer genomen?"
label var Gabapentine_use_per_day "Hoe vaak op een dag heeft u in de afgelopen 6 maanden Gabapentine genomen?"
label var Gabapentine_days_per_6months "Op hoeveel dagen in de afgelopen 6 maanden heeft u Gabapentine gebruikt?"
label var SEH_yesno "Bent u in de afgelopen 6 maanden op de spoedeisende eerste hulp van een ziekenhuis geweest?"
label var visits_SEH "Hoe vaak bent u in de afgelopen 6 maanden op de spoedeisende eerste hulp van een ziekenhuis geweest?"
label var ambulance_yesno "Bent u in de afgelopen 6 maanden met een ambulance naar het ziekenhuis gebracht?"
label var ambulance "Hoe vaak bent u in de afgelopen 6 maanden met een ambulance naar het ziekenhuis gebracht?"
label var hospital_appointment_yesno "Had u in de afgelopen 6 maanden een afspraak bij de polikliniek van het ziekenhuis?"
label var outpatient_treatment_yesno "Bent u in de afgelopen 6 maanden in het ziekenhuis geweest voor een behandeling?"
label var inpatient_treatment_yesno "Heeft u in de afgelopen 6 maanden weleens in het ziekenhuis gelegen?"
label var hospital_overnight_times "Hoe vaak heeft u in de afgelopen 6 maanden in het ziekenhuis gelegen?"
label var hospital_overnight_days "Hoe lang heeft u in de afgelopen 6 maanden in het ziekenhuis gelegen?"

label var help_due_symptoms_yesno "Heeft u in de afgelopen 6 maanden hulp gekregen van een familielid of bekende vanwege uw lichamelijke of psychische problemen?"
label var weeks_domestichelp "Hoeveel weken heeft u huishoudelijk hulp gehad?"
label var weeks_nursing "Hoeveel weken heeft u hulp gehad bij verzorging van uzelf?"
label var weeks_practicalhelp "Hoeveel weken heeft u praktische hulp gehad?"
label var hours_domestichelp "Hoeveel uur huishoudelijk hulp kreeg u in deze weken gemiddeld"
label var hours_nursing "Hoeveel uur hulp bij verzorging van uzelf kreeg u in deze weken gemiddeld"
label var hours_practicalhelp "Hoeveel uur praktische hulp kreeg u in deze weken gemiddeld"
label var transport_other_specification "Heeft u een andere wijze van vervoer gebruikt om van huis naar het ziekenhuis te gaan?"
label var distance_hospital_km "Wat was de enkele reisafstand tussen uw huis en het ziekenhuis?"

** iPCQ – In all questions the following sentence is stated: "in the past 4 weeks"

label var paid_work_yesno "Heeft u betaald werk?"
label var paid_work_hours "Hoeveel uur per week werkt u?"
label var paid_work_days "Op hoeveel dagen in de week werkt u?"
label var absenteeism_yesno "Bent u afwezig geweest van uw werk omdat u ziek was?"
label var absenteeism_days "Hoeveel werkdagen bent u afwezig geweest van uw werk omdat u ziek was?"
label var absenteeism_long "Was u langer dan de gehele periode van 4 weken afwezig van uw werk omdat u ziek was?"
label var reportingsick_day "Wanneer heeft u zich ziek gemeld (dag)?"
label var reportingsick_month "Wanneer heeft u zich ziek gemeld (maand)?"
label var reportingsick_year "Wanneer heeft u zich ziek gemeld (jaar)?"
label var presenteeism_yesno "Waren er dagen waarop u wel gewerkt heeft, maar tijdens uw werk last had van lichamelijke/psychische problemen?"
label var presenteeism_days "Op hoeveel werkdagen had u tijdens uw werk last van uw lichamelijke/psychische problemen?"
label var score_productiv_workwhensymptoms "Hoeveel werk kon u op deze dagen gemiddeld doen?"
label var missed_unpaid_work_yesno "Waren er dagen dat u minder onbetaald werk kon doen door uw lichamelijke/psychische problemen?"
label var missed_unpaid_work_days "Op hoeveel dagen kon u minder onbetaald werk doen?"
label var hours_replace_unpaidw_family "Als u hulp had gekregen op deze dagen en al het onbetaalde werk wat u niet kon doen voor u had gedaan. Hoeveel uur was die persoon hier op deze dagen dan gemiddeld mee bezig geweest?"

** EQ-5D-5L - in alle vragen staat in de vragenlijst "vandaag"

label var EQ_mobility "EQ-5D Mobiliteit"
label var EQ_selfcare "EQ-5D Zelfzorg"
label var EQ_daily_activ "EQ-5D Dagelijkse Activiteiten"
label var EQ_pain "EQ-5D Pijn / Ongemak"
label var EQ_fear "EQ-5D Angst/Somberheid"
label var EQ_VAS_score "EQ-5D VAS score, 0=worst physical health, ..., 100=best physical health"

** EHP-30

label var EHP30_skipsocials "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose niet naar sociale gelegenheden kunnen gaan door de pijn?"
label var EHP30_nohousehold_act "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose geen werkzaamheden in en/of om het huis kunnen uitvoeren door de pijn?"
label var EHP30_pain_standing "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose pijn gehad bij het staan?"
label var EHP30_pain_sitting "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose pijn gehad bij het zitten?"
label var EHP30_pain_walking "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose pijn gehad bij het lopen?"
label var EHP30_no_sports "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het moeilijk gevonden om te sporten of uw favoriete vrijetijdsbesteding uit te oefenen?"
label var EHP30_no_appetite "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose uw eetlust verloren en/of niet kunnen eten door de pijn?"
label var EHP30_sleepless "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose niet goed kunnen slapen door de pijn?"
label var EHP30_bed_due_pain "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose naar bed moeten gaan/moeten gaan liggen door de pijn?"
label var EHP30_unable_dowhatwant "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose niet de dingen kunnen doen die u wilde doen door de pijn?"
label var EHP30_unbearable_pain "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het gevoel gehad dat u de pijn niet kon verdragen?"
label var EHP30_notfeelingwell "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose uzelf in het algemeen niet lekker gevoeld?"

label var EHP30_frustrated_no_improvements "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich gefrustreerd gevoeld omdat uw symptomen niet verbeteren?"

label var EHP30_frustrated_no_control "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich gefrustreerd gevoeld omdat u uw symptomen niet kunt beheersen?"

label var EHP30_unforgettable_symptoms "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich niet in staat gevoeld om uw klachten te vergeten?"

label var EHP30_symptoms_control_life "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het gevoel gehad dat uw klachten uw leven beheersen?"

label var EHP30_symptoms_hinder_life "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het gevoel gehad dat uw klachten u beletten om te leven?"

label var EHP30_depressed "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich gedeprimeerd gevoeld?"

label var EHP30_sad "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich huilerig / verdrietig gevoeld?"

label var EHP30_miserable "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich ellendig gevoeld?"

label var EHP30_moodswings "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose stemmingswisselingen gehad?"

label var EHP30_moody "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich humeurig of kortaf gevoeld?"

label var EHP30_aggressive "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich gewelddadig of agressief gevoeld?"

label var EHP30_unable_tell_feelings "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich niet in staat gevoeld om mensen te vertellen hoe u zich voelde?"

label var EHP30_felt_misunderstood "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het gevoel gehad dat anderen niet begrijpen wat u doormaakt?"

label var EHP30_felt_whiner "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het gevoel gehad dat anderen denken dat u zeurt?"

label var EHP30_felt_lonely "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich eenzaam gevoeld?"

label var EHP30_not_desired_clothes "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose zich gefrustreerd gevoeld omdat u niet de kleren kon dragen die u zou willen?"

label var EHP30_influenced_appearance "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose het gevoel gehad dat uw uiterlijk wordt beïnvloed?"

label var EHP30_no_confidence "Gedurende de laatste 4 weken, hoe vaak heeft u door endometriose niet voldoende zelfvertrouwen gehad?"

** MHQoL

label var MHQoL_selfimage "MHQoL_ZELFBEELD"

label var MHQoL_independence "MHQoL_ONAFHANKELIJKHEID"

label var MHQoL_mood "MHQoL_STEMMING"

label var MHQoL_relationships "MHQoL_RELATIES"

label var MHQoL_dailyactiv "MHQoL_DAGELIJKSE ACTIVITEITEN"

label var MHQoL_physhealth "MHQoL_LICHAMELIJKE GEZONDHEID"

label var MHQoL_future "MHQoL_TOEKOMST"

label var MHQoL_psychological_wellbeing "MHQoL VAS, 1= worst mental well being, ..., 10=best mental well being"

** define variables

** general

label define written_consent 1 "yes" 2 "no"

label define age18plus 1 "yes" 0 "no"

label define clinical_diagnosis 1 "yes" 2 "no" 3 "don't know"

label define grade_endo 1 "Graad 1" 2 "Graad 2" 3 "Graad 3" 4 "Graad 4" 5 "Weet ik niet"

label define area_peritoneum 1 "yes" 0 "no"

label define area_bladder 1 "yes" 0 "no"

label define area_colon 1 "yes" 0 "no"

label define area_ovaries 1 "yes" 0 "no"

label define area_adenomyose 1 "yes" 0 "no"

label define area_vagina 1 "yes" 0 "no"

label define area_ureter 1 "yes" 0 "no"

label define area_unknown 1 "yes" 0 "no"

label define area_other 1 "yes" 0 "no"

label define comorb_none 1 "yes" 0 "no"

label define comorb_infertility 1 "yes" 0 "no"

label define comorb_migraine 1 "yes" 0 "no"

label define comorb_depression 1 "yes" 0 "no"

label define comorb_backpain 1 "yes" 0 "no"

label define comorb_fibroiduterus 1 "yes" 0 "no"

label define comorb_asthma 1 "yes" 0 "no"
label define comorb_thyroid 1 "yes" 0 "no"
label define comorb_fatigue 1 "yes" 0 "no"
label define comorb_PCOS 1 "yes" 0 "no"
label define comorb_other 1 "yes" 0 "no"
label define age_cat 1 "19-29 jaar" 2 "30-39 jaar" 3 "40-49 jaar" 4 "50-59 jaar"
label define marital_status 1 "Ongehuwd" 2 "Gehuwd" 3 "Partnerschap" 4 "Gescheiden" 5 "Weduwe"
label define educ_level 1 "Ik heb geen school of opleiding afgemaakt" 2 "Lagere school of basisschool" 3
"Huishoudschool, vbo, lbo, lts, leao of lhno" 4 "Mavo, mulo, ivo of vmbo" 5 "Mbo, mts, meao, inas of intas" 6 "Havo,
vwo, hbs, mms, atheneum of gymnasium" 7 "hbo, hts, heao of hhno" 8 "universiteit" 9 "Ik heb een andere opleiding
afgemaakt"
label define occupation 1 "Ik zit op school, ik studeer" 2 "Ik werk in loondienst" 3 "Ik ben zelfstandig ondernemer" 4 "Ik
ben huisvrouw" 5 "Ik ben werkloos" 6 "Ik ben arbeidsongeschikt" 7 "Ik ben met pensioen of prepensioen" 8 "Ik doe iets
anders, namelijk"

** iMCQ

label define GP_GPnurse_yesno 1 "yes" 0 "no"
label define social_worker_visits_yesno 1 "yes" 0 "no"
label define phys_therap_yesno 1 "yes" 0 "no"
label define dietician_yesno 1 "yes" 0 "no"
label define homeopath_yesno 1 "yes" 0 "no"
label define psychologist_yesno 1 "yes" 0 "no"
label define occupational_phys_yesno 1 "yes" 0 "no"
label define medicine_yesno 1 "yes" 0 "no"
label define SEH_yesno 1 "yes" 0 "no"
label define ambulance_yesno 1 "yes" 0 "no"
label define hospital_appointment_yesno 1 "yes" 0 "no"
label define outpatient_treatment_yesno 1 "yes" 0 "no"
label define inpatient_treatment_yesno 1 "yes" 0 "no"
label define help_due_symptoms_yesno 1 "yes" 0 "no"

** iPCQ

label define paid_work_yesno 1 "yes" 0 "no"
label define absenteism_yesno 1 "yes" 0 "no"
label define presenteeism_yesno 1 "yes" 0 "no"
label define score_productiv_workwhensymptoms 0 "ik kon niets doen" 10 "ik kon net zoveel doen als normaal"
label define missed_unpaid_work_yesno 1 "yes" 0 "no"

** EQ-5D-5L

label define EQ_mobility 1 "Ik heb geen problemen met lopen" 2 "Ik heb een beetje problemen met lopen" 3 "Ik heb
matige problemen met lopen" 4 "Ik heb ernstige problemen met lopen" 5 "Ik ben niet in staat om te lopen"
label define EQ_selfcare 1 "Ik heb geen problemen met mijzelf wassen of aankleden" 2 "Ik heb een beetje problemen met
mijzelf wassen of aankleden" 3 "Ik heb matige problemen met mijzelf wassen of aankleden" 4 "Ik heb ernstige problemen
met mijzelf wassen of aankleden" 5 "Ik ben niet in staat mijzelf te wassen of aan te kleden"
label define EQ_daily_activ 1 "Ik heb geen problemen met mijn dagelijkse activiteiten" 2 "Ik heb een beetje problemen
met mijn dagelijkse activiteiten" 3 "Ik heb matige problemen met mijn dagelijkse activiteiten" 4 "Ik heb ernstige
problemen met mijn dagelijkse activiteiten" 5 "Ik ben niet in staat mijn dagelijkse activiteiten uit te voeren"
label define EQ_pain 1 "Ik heb geen pijn of ongemak" 2 "Ik heb een beetje pijn of ongemak" 3 "Ik heb matige pijn of
ongemak" 4 "Ik heb ernstige pijn of ongemak" 5 "Ik heb extreme pijn of ongemak"
label define EQ_fear 1 "Ik ben niet angstig of somber" 2 "Ik ben een beetje angstig of somber" 3 "Ik ben matig angstig of
somber" 4 "Ik ben erg angstig of somber" 5 "Ik ben extreem angstig of somber"

** EHP-30

label define EHP30_skipsocials 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_nohousehold_act 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_pain_standing 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_pain_sitting 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_pain_walking 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_no_sports 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_no_appetite 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_sleepless 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_bed_due_pain 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_unable_dowhatwant 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_unbearable_pain 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_notfeelingwell 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_frustrated_no_improvements 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_frustrated_no_control 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"

```

label define EHP30_unforgettable_symptoms 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_symptoms_control_life 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_symptoms_hinder_life 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_depressed 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_sad 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_miserable 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_moodswings 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_moody 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_aggressive 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_unable_tell_feelings 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_felt_misunderstood 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_felt_whiner 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_felt_lonely 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_not_desired_clothes 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_influenced_appearance 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"
label define EHP30_no_confidence 1 "Nooit" 2 "Zelden" 3 "Soms" 4 "Vaak" 5 "Altijd"

```

```
** MHQoL
```

```

label define MHQoL_selfimage 1 "Ik denk heel positief over mijzelf" 2 "Ik denk positief over mijzelf" 3 "Ik denk negatief over mijzelf" 4 "Ik denk heel negatief over mijzelf"
label define MHQoL_independence 1 "Ik ben heel tevreden over de mate van mijn onafhankelijkheid" 2 "Ik ben tevreden over de mate van mijn onafhankelijkheid" 3 "Ik ben ontevreden over de mate van mijn onafhankelijkheid" 4 "Ik ben heel ontevreden over de mate van mijn onafhankelijkheid"
label define MHQoL_mood 1 "Ik voel mij niet angstig, somber of depressief" 2 "Ik voel mij een beetje angstig, somber of depressief" 3 "Ik voel mij angstig, somber of depressief" 4 "Ik voel mij heel angstig, somber of depressief"
label define MHQoL_relationships 1 "Ik ben heel tevreden over mijn relaties" 2 "Ik ben tevreden over mijn relaties" 3 "Ik ben ontevreden over mijn relaties" 4 "Ik ben heel ontevreden over mijn relaties"
label define MHQoL_dailyactiv 1 "Ik ben heel tevreden over mijn dagelijkse activiteiten" 2 "Ik ben tevreden over mijn dagelijkse activiteiten" 3 "Ik ben ontevreden over mijn dagelijkse activiteiten" 4 "Ik ben heel ontevreden over mijn dagelijkse activiteiten"
label define MHQoL_physhealth 1 "Ik heb geen probleem met mijn lichamelijke gezondheid" 2 "Ik heb enkele problemen met mijn lichamelijke gezondheid" 3 "Ik heb veel problemen met mijn lichamelijke gezondheid" 4 "Ik heb zeer veel problemen met mijn lichamelijke gezondheid"
label define MHQoL_future 1 "Ik zie mijn toekomst heel positief in" 2 "Ik zie mijn toekomst positief in" 3 "Ik zie mijn toekomst somber in" 4 "Ik zie mijn toekomst heel somber in"

```

```
///// changes GENERAL QUESTIONS /////
```

```

replace age18plus=1 if age18plus==1
replace age18plus=0 if age18plus!=1

```

```
** Manually change the "other."-option in Stata for area_other_specification
```

```

replace area_peritoneum = 1 if ID == 73 | ID == 120
replace area_bladder = 1 if ID == 9 | ID == 50
replace area_colon = 1 if ID == 147 | ID == 173 | ID == 281 | ID == 241
replace area_ovaries = 1 if ID == 173 | ID == 271 | ID == 277 | ID == 314 | ID == 126
replace area_adenomyose = 1 if ID == 168
replace area_vagina = 1 if ID == 49

```

```
** Change/add grouping of areas
```

```

gen area_ligaments = 0
label var area_ligaments "Is er endometriose aangetroffen op/in de ligamenten uterus en torus uteri?"
label define area_ligaments 1 "yes" 2 "no"
order area_ligaments, a(area_ureter)
replace area_ligaments = 1 if ID == 305 | ID == 181 | ID == 44 | ID == 73 | ID == 96 | ID == 100 | ID == 173 | ID == 243 | ID == 215 | ID == 20 | ID == 205 | ID == 271 | ID == 241 | ID == 70 | ID == 32 | ID == 67 | ID == 293 | ID == 94 | ID == 315 | ID == 85 | ID == 187 | ID == 142 | ID == 83 | ID == 4 | ID == 302

```

```
gen area_thoracic = 0
```

```
label var area_thoracic "Is er thoracale endometriose aangetroffen (middenrif, thorax, hartvlies)"
```

```

label define area_thoracic 1 "yes" 0 "no"
order area_thoracic, a(area_ligaments)
replace area_thoracic = 1 if ID == 73 | ID == 173 | ID == 27 | ID == 44 | ID == 236 | ID == 277 | ID == 161 | ID == 120

```

```

replace area_other = 0 if ID == 181 | ID == 49 | ID == 168 | ID == 9 | ID == 50 | ID == 258 | ID == 314 | ID == 34 | ID == 281 | ID == 73 | ID == 70 | ID == 271 | ID == 67 | ID == 293 | ID == 94 | ID == 315 | ID == 85 | ID == 193 | ID == 27 | ID

```

```
== 305 | ID == 44 | ID == 142 | ID == 236 | ID == 83 | ID == 4 | ID == 302 | ID == 277 | ID == 120 | ID == 100 | ID ==
96 | ID == 243 | ID == 215 | ID == 20 | ID == 126 | ID == 241
```

```
** Manually change the "other."-option in Stata for comorbidities_otherspecification
replace comorb_migraine=1 if ID==191
replace comorb_depression=1 if ID==126 | ID == 210 | ID == 193 | ID == 207
replace comorb_fibroiduterus=1 if ID == 314
replace comorb_fatigue=1 if ID==126
replace comorb_other = 0 if ID == 210 | ID == 314 | ID == 126
```

```
** Manually change the "other."-option in Stata for educ_level_other
replace educ_level=7 if ID == 173
```

```
** Manually change the "other."-option in Stata for occupation_other
replace occupation = 2 if ID == 11
```

```
** DROP
```

```
* drop educ_level_other because none present anymore
drop educ_level_other
```

```
* not having finished full survey --> 105 observations removed
drop if Progress < 98
```

```
* not giving written consent --> 25 observations removed.
drop if written_consent != 1
```

```
* not being 18 or above (adults) --> 1 observations removed.
drop if age18plus != 1
```

```
** not getting a clinical diagnosis via MRI, echocardiogram or laparoscopy (or didn't know if they got a diagnosis) = 2
and 3. --> 3 observations deleted.
drop if clinical_diagnosis != 1
```

```
** Not knowing which grade endometriosis is detected --> 53 observations deleted.
drop if grade_endo == 5
```

```
** to correct for mistakes in first stage of survey --> 33 observations deleted
drop if GP_GPnurse_yesno == .
```

```
** Year of symptoms cannot be missing --> 5 observations deleted
drop if year_symptoms == .
```

```
** 0 observations deleted
drop if diagnostic_delay == .
```

```
** diagnostic delay cannot be below 0 --> 1 observations deleted
drop if diagnostic_delay < 0
```

```
** diagnostic delay cannot be larger than a lifetime --> 1 observations deleted
drop if diagnostic_delay > 100
```

```
drop year_symptoms
drop year_diagnosis
drop written_consent
```

```
///// changes general questionnaire \\\\\\\
```

```
** create dummy for when there are comorbidities or there are not
```

```
gen comorb_present = .
```

```
replace comorb_present = 1 if comorb_infertility == 1 | comorb_migraine == 1 | comorb_depression == 1 |
comorb_backpain == 1 | comorb_fibroiduterus == 1 | comorb_asthma == 1 | comorb_thyroid == 1 | comorb_fatigue == 1 |
comorb_PCOS == 1 | comorb_other == 1 | comorb_none == 0
```

```
replace comorb_present = 0 if comorb_none == 1
```

```
order comorb_present, a(comorb_other)
```

```
label var comorb_present "Any comorbidity being present"
```

```
label define comorb_present 1 "yes" 0 "no"
```

```

///// changes iMCQ \\\\\ 6 months
replace GP_GPnurse_yesno = 0 if GP_GPnurse_yesno == 1
replace GP_GPnurse_yesno = 1 if GP_GPnurse_yesno == 2
drop GP_yes_POH_yes

** For achieving valid averages, filled in zero-values are replaced by missing values.
replace GP_visits = 0 if GP_visits == .
replace POH_visits = 0 if POH_visits == .

replace social_worker_visits_yesno = 0 if social_worker_visits_yesno == 1
replace social_worker_visits_yesno = 1 if social_worker_visits_yesno == 2
replace social_worker_visits = 0 if social_worker_visits == .

replace phys_therap_yesno = 0 if phys_therap_yesno == 1
replace phys_therap_yesno = 1 if phys_therap_yesno == 2
replace phys_therap_visits = 1 if phys_therap_visits == .

replace dietician_yesno = 0 if dietician_yesno == 1
replace dietician_yesno = 1 if dietician_yesno == 2
replace dietician_visits = 0 if dietician_visits == .

replace homeopath_yesno = 0 if homeopath_yesno == 1
replace homeopath_yesno = 1 if homeopath_yesno == 2
replace homeopath_visits = 0 if homeopath_visits == .

replace psychologist_yesno = 0 if psychologist_yesno == 1
replace psychologist_yesno = 1 if psychologist_yesno == 2
replace psychologist_visits = 0 if psychologist_visits == .

replace occupational_phys_yesno = 0 if occupational_phys_yesno == 1
replace occupational_phys_yesno = 1 if occupational_phys_yesno == 2
replace occupational_phys_visits = 0 if occupational_phys_visits == .

** Medicine use
** PAIN MEDICATION
replace medicine_yesno = 0 if medicine_yesno == 1
replace medicine_yesno = 1 if medicine_yesno == 2

replace paracetamol_per_use=0 if missing(paracetamol_per_use)
replace paracetamol_use_per_day=0 if missing(paracetamol_use_per_day)
replace paracetamol_days_per_6months=0 if missing(paracetamol_days_per_6months)

gen paracetamol_yesno = 0
replace paracetamol_yesno = 1 if paracetamol_per_use != 0 | paracetamol_use_per_day != 0 |
paracetamol_days_per_6months != 0
label var paracetamol_yesno "Whether or not a patient used paracetamol in the past 6 months"
label define paracetamol_yesno 1 "yes" 0 "no"
order paracetamol_yesno, a(paracetamol_days_per_6months)

replace NSAID_per_use=0 if missing(NSAID_per_use)
replace NSAID_use_per_day=0 if missing(NSAID_use_per_day)
replace NSAID_days_per_6months=0 if missing(NSAID_days_per_6months)

gen NSAID_yesno = 0
replace NSAID_yesno = 1 if NSAID_per_use != 0 | NSAID_use_per_day != 0 | NSAID_days_per_6months != 0
label var NSAID_yesno "Whether or not a patient used NSAIDs in the past 6 months"
label define NSAID_yesno 1 "yes" 0 "no"
order NSAID_yesno, a(NSAID_days_per_6months)

replace Tramadol_per_use=0 if missing(Tramadol_per_use)
replace Tramadol_use_per_day=0 if missing(Tramadol_use_per_day)
replace Tramadol_days_per_6months=0 if missing(Tramadol_days_per_6months)

gen Tramadol_yesno = 0
replace Tramadol_yesno = 1 if Tramadol_per_use != 0 | Tramadol_use_per_day != 0 | Tramadol_days_per_6months != 0

```

```

label var Tramadol_yesno "Whether or not a patient used Tramadol in the past 6 months"
label define Tramadol_yesno 1 "yes" 0 "no"
order Tramadol_yesno, a(Tramadol_days_per_6months)

replace Oxynorm_per_use=0 if missing(Oxynorm_per_use)
replace Oxynorm_use_per_day=0 if missing(Oxynorm_use_per_day)
replace Oxynorm_days_per_6months=0 if missing(Oxynorm_days_per_6months)

gen Oxynorm_yesno = 0
replace Oxynorm_yesno = 1 if Oxynorm_per_use != 0 | Oxynorm_use_per_day != 0 | Oxynorm_days_per_6months != 0
label var Oxynorm_yesno "Whether or not a patient used Oxynorm in the past 6 months"
label define Oxynorm_yesno 1 "yes" 0 "no"
order Oxynorm_yesno, a(Oxynorm_days_per_6months)

replace Oxycontin_per_use=0 if missing(Oxycontin_per_use)
replace Oxycontin_use_per_day=0 if missing(Oxycontin_use_per_day)
replace Oxycontin_days_per_6months=0 if missing(Oxycontin_days_per_6months)

gen Oxycontin_yesno = 0
replace Oxycontin_yesno = 1 if Oxycontin_per_use != 0 | Oxycontin_use_per_day != 0 | Oxycontin_days_per_6months != 0
label var Oxycontin_yesno "Whether or not a patient used Oxycontin in the past 6 months"
label define Oxycontin_yesno 1 "yes" 0 "no"
order Oxycontin_yesno, a(Oxycontin_days_per_6months)

replace Gabapentine_per_use=0 if missing(Gabapentine_per_use)
replace Gabapentine_use_per_day=0 if missing(Gabapentine_use_per_day)
replace Gabapentine_days_per_6months=0 if missing(Gabapentine_days_per_6months)

gen Gabapentine_yesno = 0
replace Gabapentine_yesno = 1 if Gabapentine_per_use != 0 | Gabapentine_use_per_day != 0 | Gabapentine_days_per_6months != 0
label var Gabapentine_yesno "Whether or not a patient used Gabapentine in the past 6 months"
label define Gabapentine_yesno 1 "yes" 0 "no"
order Gabapentine_yesno, a(Gabapentine_days_per_6months)

** OTHER PAIN MEDICATION
** Manually change the "other.."option in Stata for other pain medication
replace NSAID_per_use = 1 if ID == 242
replace NSAID_use_per_day = 2 if ID == 242
replace NSAID_days_per_6months = 182 if ID == 242

* drop CF | CG | CH | CI

** HORMONAL THERAPY
destring oral_contraceptive, replace
replace oral_contraceptive = 0 if missing(oral_contraceptive)
label var oral_contraceptive "Heeft u in de afgelopen 6 maanden gebruik gemaakt van orale anticonceptie?"
label define oral_contraceptive 1 "yes" 0 "no"

destring hormonal_IUD, replace
replace hormonal_IUD=0 if missing(hormonal_IUD)
label var hormonal_IUD "Heeft u in de afgelopen 6 maanden gebruik gemaakt van een hormoonhoudend spiraal?"
label define hormonal_IUD 1 "yes" 0 "no"

destring oral_progestativa, replace
replace oral_progestativa=0 if missing(oral_progestativa)
label var oral_progestativa "Heeft u in de afgelopen 6 maanden gebruik gemaakt van een orale progestativa?"
label define oral_progestativa 1 "yes" 0 "no"

destring GNRH_analoog, replace
replace GNRH_analoog=0 if missing(GNRH_analoog)
label var GNRH_analoog "Heeft u in de afgelopen 6 maanden gebruik gemaakt van een GNRH analoog?"
label define GNRH_analoog 1 "yes" 0 "no"

destring hormonal_ther_other, replace

```

```
replace hormonal_ther_other=0 if missing(hormonal_ther_other)
label var hormonal_ther_other "Heeft u in de afgelopen 6 maanden gebruik gemaakt van een andere hormonale therapie?"
label define hormonal_ther_other 1 "yes" 0 "no"
```

```
** OTHER IN GENERAL
```

```
ren Overignamelijkttekst medicine_other_overall
ren OverignamelijkHoeveelheef medicine_other_overall_per_use
replace medicine_other_overall_per_use=0 if missing(medicine_other_overall_per_use)
ren OverignamelijkHoevaakop medicine_other_overallper_day
replace medicine_other_overallper_day=0 if missing(medicine_other_overallper_day)
ren OverignamelijkOphoeveeld medicine_other_overall6months
replace medicine_other_overall6months=0 if missing(medicine_other_overall6months)
```

```
** SEH CARE
```

```
replace SEH_yesno = 0 if SEH_yesno == 1
replace SEH_yesno = 1 if SEH_yesno == 2
replace visits_SEH = 0 if visits_SEH == .
```

```
** AMBULANCE
```

```
replace ambulance_yesno = 0 if ambulance_yesno == 1
replace ambulance_yesno = 1 if ambulance_yesno == 2
destring ambulance, replace
replace ambulance = 0 if ambulance == .
```

```
** PHYSICIAN APPOINTMENT HOSPITAL
```

```
replace hospital_appointment_yesno = 0 if hospital_appointment_yesno == 1
replace hospital_appointment_yesno = 1 if hospital_appointment_yesno == 2
drop CY
```

```
** GYNAECOLOGIST
```

```
gen gynaecologist = .
order gynaecologist, a(hospital_appointment_yesno)
replace gynaecologist = 1 if ID == 129 | ID == 203 | ID == 215 | ID == 174 | ID == 287 | ID == 278 | ID == 36 | ID ==
134 | ID == 113 | ID == 259 | ID == 276 | ID == 245 | ID == 239 | ID == 138 | ID == 275 | ID == 37 | ID == 319 | ID ==
100 | ID == 141 | ID == 43 | ID == 170 | ID == 248 | ID == 50 | ID == 126 | ID == 220 | ID == 214 | ID == 299
replace gynaecologist = 2 if ID == 92 | ID == 190 | ID == 216 | ID == 27 | ID == 222 | ID == 44 | ID == 236 | ID == 169 |
ID == 106 | ID == 66 | ID == 151 | ID == 61 | ID == 149 | ID == 260 | ID == 146 | ID == 79 | ID == 148 | ID == 9 | ID ==
135 | ID == 70 | ID == 280 | ID == 315 | ID == 193 | ID == 243 | ID == 1
replace gynaecologist = 3 if ID == 234 | ID == 48 | ID == 142 | ID == 83 | ID == 296 | ID == 300 | ID == 12 | ID == 118 |
ID == 31 | ID == 14 | ID == 39 | ID == 111 | ID == 262 | ID == 241 | ID == 183 | ID == 250 | ID == 242 | ID == 67 | ID
== 232 | ID == 207 | ID == 85
replace gynaecologist = 4 if ID == 130 | ID == 191 | ID == 305 | ID == 205 | ID == 212 | ID == 208 | ID == 161 | ID ==
304 | ID == 73 | ID == 96 | ID == 168 | ID == 179 | ID == 228 | ID == 16 | ID == 103 | ID == 221 | ID == 230 | ID == 181
| ID == 271 | ID == 81 | ID == 303 | ID == 293
replace gynaecologist = 5 if ID == 187 | ID == 121 | ID == 8 | ID == 261 | ID == 29 | ID == 257
replace gynaecologist = 6 if ID == 17 | ID == 7 | ID == 238 | ID == 32 | ID == 273
replace gynaecologist = 7 if ID == 140 | ID == 10 | ID == 264 | ID == 6
replace gynaecologist = 8 if ID == 253
replace gynaecologist = 9 if ID == 206
replace gynaecologist = 20 if ID == 153
replace gynaecologist = 35 if ID == 94
label var gynaecologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de gynaecoloog gehad?"
replace gynaecologist = 0 if gynaecologist == .
```

```
** INTERNIST
```

```
gen internist = .
order internist, a(gynaecologist)
replace internist = 1 if ID == 121
replace internist = 2 if ID == 284
replace internist = 4 if ID == 66
replace internist = 5 if ID == 253 | ID == 315
label var internist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de internist gehad?"
replace internist = 0 if internist == .
```

```
** UROLOGIST
```

```
gen urologist = .
order urologist, a(internist)
```

```

replace urologist = 1 if ID == 262 | ID == 120 | ID == 122
replace urologist = 2 if ID == 121 | ID == 208 | ID == 168
replace urologist = 3 if ID == 191
replace urologist = 6 if ID == 111
replace urologist = 8 if ID == 103
label var urologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de uroloog gehad?"
replace urologist = 0 if urologist == .

** NEUROLOGIST
gen neurologist = .
order neurologist, a(urologist)
replace neurologist = 1 if ID == 16
replace neurologist = 2 if ID == 191 | ID == 32
replace neurologist = 3 if ID == 230
label var neurologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de neuroloog gehad?"
replace neurologist = 0 if neurologist == .

** GASTRO-ENTEROLOGIST
gen gastro_enterologist = .
order gastro_enterologist, a(neurologist)
replace gastro_enterologist = 1 if ID == 214 | ID == 79 | ID == 264 | ID == 193 | ID == 6
replace gastro_enterologist = 2 if ID == 130 | ID == 27 | ID == 49 | ID == 134 | ID == 126
replace gastro_enterologist = 3 if ID == 299 | ID == 53 | ID == 135 | ID == 103
replace gastro_enterologist = 4 if ID == 253 | ID == 61 | ID == 257
replace gastro_enterologist = 5 if ID == 220 | ID == 303
label var gastro_enterologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de MDL-arts gehad?"
replace gastro_enterologist = 0 if gastro_enterologist == .

** ANAESTHESIST (OOK PIJNPOLI)
gen anaesthesist = .
order anaesthesist, a(gastro_enter)
replace anaesthesist = 1 if ID == 129 | ID == 278 | ID == 244 | ID == 161 | ID == 250 | ID == 181 | ID == 200
replace anaesthesist = 2 if ID == 206 | ID == 6
replace anaesthesist = 3 if ID == 191 | ID == 103
replace anaesthesist = 4 if ID == 44
label var anaesthesist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de anesthesist gehad?"
replace anaesthesist = 0 if anaesthesist == .

** RADIOLOGIST
gen radiologist = .
order radiologist, a(anaesthesist)
replace radiologist = 1 if ID == 12 | ID == 151 | ID == 229 | ID == 70 | ID == 228 | ID == 191
replace radiologist = 2 if ID == 140 | ID == 264 | ID == 37
label var radiologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de radioloog gehad?"
replace radiologist = 0 if radiologist == .

** PULMONOLOGIST
gen pulmonologist = .
order pulmonologist, a(radiologist)
replace pulmonologist = 1 if ID == 236 | ID == 6
replace pulmonologist = 2 if ID == 118 | ID == 138
label var pulmonologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de longarts gehad?"
replace pulmonologist = 0 if pulmonologist == .

** ULTRASOUND SPECIALIST
gen ultrasound_specialist = .
order ultrasound_specialist, a(pulmonologist)
replace ultrasound_specialist = 1 if ID == 303 | ID == 228
label var ultrasound_specialist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de echoscopist gehad?"
replace ultrasound_specialist = 0 if ultrasound_specialist == .

** SURGEON
gen surgeon = .
order surgeon, a(ultrasound_specialist)
replace surgeon = 1 if ID == 244 | ID == 146 | ID == 168 | ID == 39 | ID == 238 | ID == 71 | ID == 103 | ID == 250
replace surgeon = 2 if ID == 130 | ID == 305 | ID == 79 | ID == 153
replace surgeon = 3 if ID == 303

```

```

replace surgeon = 4 if ID == 10
replace surgeon = 5 if ID == 73
label var surgeon "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de chirurg gehad?"
replace surgeon = 0 if surgeon == .

** PSYCHIATRIST
gen psychiatrist = .
order psychiatrist, a(surgeon)
replace psychiatrist = 3 if ID == 243 | ID == 304
label var psychiatrist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de psychiater gehad?"
replace psychiatrist = 0 if psychiatrist == .

** FERTILITY SPECIALIST
gen fertility_specialist = .
order fertility_specialist, a(psychiatrist)
replace fertility_specialist = 1 if ID == 142
replace fertility_specialist = 2 if ID == 10
replace fertility_specialist = 4 if ID == 4
replace fertility_specialist = 9 if ID == 315
replace fertility_specialist = 10 if ID == 299 | ID == 161 | ID == 14
replace fertility_specialist = 12 if ID == 58
replace fertility_specialist = 15 if ID == 245
replace fertility_specialist = 17 if ID == 261
replace fertility_specialist = 20 if ID == 179
label var fertility_specialist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de fertilititeitsarts gehad?"
replace fertility_specialist = 0 if fertility_specialist == .

** DERMATOLOGIST
gen dermatologist = .
order dermatologist, a(fertility_specialist)
replace dermatologist = 1 if ID == 254
replace dermatologist = 2 if ID == 205 | ID == 302
label var dermatologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de dermatoloog gehad?"
replace dermatologist = 0 if dermatologist == .

** SEXOLOGIST
gen sexologist = .
order sexologist, a(dermatologist)
replace sexologist = 2 if ID == 244
label var sexologist "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij de seksuoloog gehad?"
replace sexologist = 0 if sexologist == .

** OTHER - general
gen spec_visit_other = .
order spec_visit_other, a(sexologist)
replace spec_visit_other = 1 if ID == 4
label var spec_visit_other "Hoe vaak heeft u in de afgelopen 6 maanden een afspraak bij een andere dokter in het ziekenhuis gehad?"
replace spec_visit_other = 0 if spec_visit_other == .

drop hospital_appointment_yesno Bijwelkesoortendoktersbentu DA DB DC DD DE DF DG DH DI DJ DK DL DM

** NOTE: some filled in specialist visits were not linked to endometriosis at all (discussed with endometriosis specialist Noortje). Those were excluded from the visits. The ones meant here are cardiologist, rheumatologist, oncologist, ENT specialist, hematologist and immunologist
egen total_specialists_visits_pp = rowtotal(gynaecologist internist urologist neurologist gastro_enterologist anaesthesist radiologist pulmonologist ultrasound_specialist surgeon psychiatrist fertility_specialist dermatologist sexologist spec_visit_other)
label var total_specialists_visits_pp "Total specialists visits (summed all specialists visits)"
order total_specialists_visits_pp, a(spec_visit_other)

egen total_other_spec_visits_pp = rowtotal(internist urologist neurologist gastro_enterologist anaesthesist radiologist pulmonologist ultrasound_specialist surgeon psychiatrist dermatologist sexologist spec_visit_other)
label var total_other_spec_visits_pp "Total other specialists visits (summed all specialists visits excluding gynaecologist and fertility specialist visits)"
order total_other_spec_visits_pp, a(total_specialists_visits_pp)

```

```

** OUTPATIENT TREATMENT = DAGBEHANDELING
replace outpatient_treatment_yesno = 0 if outpatient_treatment_yesno == 1
replace outpatient_treatment_yesno = 1 if outpatient_treatment_yesno == 2

** MRI
gen mri = .
label var mri "Number of MRI scans in the past 6 months"
order mri, a(outpatient_treatment_yesno)
replace mri = 1 if ID == 130 | ID == 10 | ID == 16
replace mri = 2 if ID == 7
replace mri = 0 if mri == .

** X-RAY = CT SCAN/RONTGEN
gen xray = .
label var xray "Number of x-rays in the past 6 months"
order xray, a(mri)
replace xray = 1 if ID == 191 | ID == 16
replace xray = 2 if ID == 61
replace xray = 0 if xray == .

** ULTRASOUND SCAN = ECHO
gen ultrasound = .
label var ultrasound "Number of ultrasound scans in the past 6 months"
order ultrasound, a(xray)
replace ultrasound = 1 if ID == 7 | ID == 120 | ID == 303
replace ultrasound = 0 if ultrasound == .

** LAPOROSCOPY = kijkoperatie
gen laparoscopy = .
label var laparoscopy "Number of laparoscopy treatments in the past 6 months"
order laparoscopy, a(ultrasound)
replace laparoscopy = 1 if ID == 142 | ID == 4 | ID == 253 | ID == 179 | ID == 181 | ID == 210
replace laparoscopy = 2 if ID == 187 | ID == 293
replace laparoscopy = 3 if ID == 121 | ID == 32
replace laparoscopy = 0 if laparoscopy == .

** ENDOSCOPYY (coloscopy (= colon), cystoscopy (= bladder), hysteroscopy (= uterus), gastroscopy)
gen endoscopy = .
label var endoscopy "Number of endoscopy treatments in the past 6 months (including coloscopy, cystoscopy,
hysteroscopy, gastroscopy)"
order endoscopy, a(laparoscopy)
replace endoscopy = 1 if ID == 130 | ID == 53 | ID == 118 | ID == 191 | ID == 121
replace endoscopy = 2 if ID == 135
replace endoscopy = 3 if ID == 257
replace endoscopy = 5 if ID == 303 | ID == 61
replace endoscopy = 0 if endoscopy == .

** SURGERY
gen surgery = .
label var surgery "Number of surgeries per patient in the past 6 months"
order surgery, a(endoscopy)
replace surgery = 1 if ID == 205 | ID == 10 | ID == 238 | ID == 48
replace surgery = 0 if surgery == .

** FOLLICEL PUNCTURE = EICELPUNCTIE/PUNCTIE IVF
gen ivf_puncture = .
label var ivf_puncture "Number of IVF puncture treatments per patient in the past 6 months"
order ivf_puncture, a(surgery)
replace ivf_puncture = 1 if ID == 315 | ID == 58
replace ivf_puncture = 3 if ID == 4
replace ivf_puncture = 4 if ID == 161
replace ivf_puncture = 0 if ivf_puncture == .

** EMBRYO TRANSFER
gen ivf_embryo_transfer = .
label var ivf_embryo_transfer "Number of IVF embryo transfer treatments per patient in the past 6 months"
order ivf_embryo_transfer, a(ivf_puncture)

```

```

replace ivf_embryo_transfer = 2 if ID == 58 | ID == 71
replace ivf_embryo_transfer = 3 if ID == 315
replace ivf_embryo_transfer = 4 if ID == 4
replace ivf_embryo_transfer = 10 if ID == 299
replace ivf_embryo_transfer = 0 if ivf_embryo_transfer == .

** FULL IVF
gen ivf_full = .
label var ivf_full "Number of full IVF treatments (indicated in survey) per patient in the past 6 months"
order ivf_full, a(ivf_embryo_transfer)
replace ivf_full = 4 if ID == 187
replace ivf_full = 5 if ID == 179
replace ivf_full = 6 if ID == 170
replace ivf_full = 15 if ID == 245
replace ivf_full = 24 if ID == 261
replace ivf_full = 0 if ivf_full == .

** LUCRIN DEPOT
gen lucrin_depot = .
label var lucrin_depot "Number of lucrin depot treatments per patient in the past 6 months"
order lucrin_depot, a(ivf_full)
replace lucrin_depot = 1 if ID == 187
replace lucrin_depot = 0 if lucrin_depot == .

** PULMONARY FUNCTION TEST
gen pulmonary_function = .
label var pulmonary_function "Number of pulmonary function tests per patient in the past 6 months"
order pulmonary_function, a(lucrin_depot)
replace pulmonary_function = 2 if ID == 118
replace pulmonary_function = 0 if pulmonary_function == .

** CURETTAGE
gen curettage = .
label var curettage "Number of curettage treatments per patient in the past 6 months"
order curettage, a(pulmonary_function)
replace curettage = 2 if ID == 161
replace curettage = 0 if curettage == .

** PLACEMENT DOUBLE-J STENT
gen placement_doublej = .
label var placement_doublej "Number of double-J placements per patient in the past 6 months"
order placement_doublej, a(curettage)
replace placement_doublej = 3 if ID == 103
replace placement_doublej = 0 if placement_doublej == .

** REMOVAL DOUBLE-J STENT
gen removal_doublej = .
label var removal_doublej "Number of double-J stent removals per patient in the past 6 months"
order removal_doublej, a(placement_doublej)
replace removal_doublej = 1 if ID == 262 | ID == 103
replace removal_doublej = 0 if removal_doublej == .

** IUI TREATMENT
gen iui = .
label var iui "Number of iui treatments per patient in the past 6 months"
order iui, a(removal_doublej)
replace iui = 20 if ID == 179 | ID == 153
replace iui = 0 if iui == .

** HCG TEST
gen hcg = .
label var hcg "Number of HCG tests per patient in the past 6 months"
order hcg, a(iui)
replace hcg = 1 if ID == 221
replace hcg = 0 if hcg == .

** CYST PUNCTURE

```

```

gen cyst_puncture = .
label var cyst_puncture "Number of cyst punctures per patient in the past 6 months"
order cyst_puncture, a(hcg)
replace cyst_puncture = 1 if ID == 170
replace cyst_puncture = 0 if cyst_puncture == .

** HSG ((FOTO))
gen hsg = .
label var hsg "Number hsg treatments per patient in the past 6 months"
order hsg, a(cyst_puncture)
replace hsg = 1 if ID == 5
replace hsg = 0 if hsg == .

** CORDOTOMY = zenuwbehandeling
gen cordotomy = .
label var cordotomy "Number of curdotomies per patient in the past 6 months"
order cordotomy, a(hsg)
replace cordotomy = 2 if ID == 16
replace cordotomy = 0 if cordotomy == .

** OTHER
gen treatment_other = .
label var treatment_other "Other, not above mentioned outpatient treatments"
order treatment_other, a(cordotomy)
replace treatment_other = 1 if ID == 236 | ID == 304 | ID == 7 | ID == 183
replace treatment_other = 2 if ID == 302 | ID == 61 | ID == 16 | ID == 257 | ID == 44
replace treatment_other = 0 if treatment_other == .

** Total number of outpatient treatments per patient
egen tot_outpatient_pp = rowtotal(mri xray ultrasound laparoscopy endoscopy surgery ivf_puncture ivf_embryo_transfer
ivf_full lucrin_depot pulmonary_function curettage placement_doublej removal_doublej iui hcg cyst_puncture hsg
cordotomy treatment_other)
label var tot_outpatient_pp "Total outpatient care/treatments in hospital in past 6 months per patient"
order tot_outpatient_pp, a(treatment_other)

drop Voorwelkesoortbehandelingwas DP DQ DR DS DT Hoeveelkeermoestuindeafgel DV DW DX DY DZ

** HOSPITAL OVERNIGHT
replace inpatient_treatment_yesno = 0 if inpatient_treatment_yesno == 1
replace inpatient_treatment_yesno = 1 if inpatient_treatment_yesno == 2

replace hospital_overnight_days = 0 if hospital_overnight_days == .
replace hospital_overnight_times = 0 if hospital_overnight_times == .

generate total_overnight_duration = hospital_overnight_days*hospital_overnight_times
label var total_overnight_duration "Total overnight duration of inpatient care in past 6 months"
order total_overnight_duration, a(hospital_overnight_days)

** DOMESTIC, NURSING, PRACTICAL HELP
drop Watvoorhulpvanfamilieleden

replace help_due_symptoms_yesno = 0 if help_due_symptoms_yesno == 1
replace help_due_symptoms_yesno = 1 if help_due_symptoms_yesno == 2

replace weeks_domestichelp = 0 if weeks_domestichelp == .
replace hours_domestichelp = 0 if hours_domestichelp == .
replace weeks_nursing = 0 if weeks_nursing == .
replace hours_nursing = 0 if hours_nursing == .
replace weeks_practicalhelp = 0 if weeks_practicalhelp == .
replace hours_practicalhelp = 0 if hours_practicalhelp == .

order hours_domestichelp, a(weeks_domestichelp)
order hours_nursing, a(weeks_nursing)
order hours_practicalhelp, a(weeks_practicalhelp)

generate total_domestichelp = weeks_domestichelp*hours_domestichelp
label var total_domestichelp "Total domestic help in the past 6 months due to pain per patient (hours)"

```

```

order total_domestichelp, a(hours_domestichelp)

generate total_nursing = weeks_nursing*hours_nursing
label var total_nursing "Total help with nursing in the past 6 months due to pain per patient (hours)"
order total_nursing, a(hours_nursing)

generate total_practicalhelp = weeks_practicalhelp*hours_practicalhelp
label var total_practicalhelp "Total practical help in the past 6 months due to pain per patient (hours)"
order total_practicalhelp, a(hours_practicalhelp)

** replace to maximum number of weeks in 6 months (for people that have more than that)
replace weeks_nursing = 26 if ID == 243
replace weeks_practicalhelp = 26 if ID == 173

** TRANSPORT
** drop because nobody filled it in
drop transport_other
drop transport_other_specification

///// iPCQ \\\\\ 4 weeks
** paid work - absenteeism
* when missing values, no paid work: does not mean 0 hours for all other questions.
replace paid_work_yesno = 0 if paid_work_yesno == 1
replace paid_work_yesno = 1 if paid_work_yesno == 2

gen hours_per_workday = (paid_work_hours/paid_work_days)
label var hours_per_workday "Hoeveel uur per dag verricht u betaald werk?"
order hours_per_workday, a(paid_work_days)

replace absenteeism_yesno = 0 if absenteeism_yesno == 1
replace absenteeism_yesno = 1 if absenteeism_yesno == 2

replace absenteeism_days = 0 if absenteeism_yesno == 0
replace absenteeism_long = 0 if absenteeism_long == 1
replace absenteeism_long = 1 if absenteeism_long == 2
replace absenteeism_long = 0 if absenteeism_yesno == 0

** correction for double counting absence days for women with long absence
** = 0 means having been absent, but not for the full 4 weeks or longer.
** = . means not having been absent at all or not having a paid job.
replace absenteeism_days = 0 if absenteeism_long == 1

** drop observations that say having been absent but not how long
drop if absenteeism_yesno == 1 & absenteeism_days == .

** missing values when women did fill in
gen total_absenteism_hours = (hours_per_workday * absenteeism_days)
label var total_absenteism_hours "Hoeveel uur bent u in de afgelopen 4 weken absent geweest van uw betaalde werk?"
order total_absenteism_hours, a(absenteism_days)

gen dailydate = dofc(RecordedDate)
gen long_absence_date = mdy(reportingsick_month, reportingsick_day, reportingsick_year)

gen calendardays_long_absence = dailydate - long_absence_date
label var calendardays_long_absence "Hoeveel kalenderdagen bent u in de afgelopen 4 weken absent geweest van uw betaalde werk (meer dan 4 weken)?"
order calendardays_long_absence, a(absenteism_long)

drop RecordedDate
drop dailydate
drop long_absence_date
drop reportingsick_month
drop reportingsick_day
drop reportingsick_year

```

```

** paid work - presenteeism
replace presenteeism_yesno = 0 if presenteeism_yesno==1
replace presenteeism_yesno = 1 if presenteeism_yesno==2

replace presenteeism_days = 0 if presenteeism_yesno == 0

drop Opdedagendatulasthadkon

** unpaid work
* when missing values, no unpaid work: does not mean 0 hours for all other questions.
replace missed_unpaid_work_yesno = 0 if missed_unpaid_work_yesno == 1
replace missed_unpaid_work_yesno = 1 if missed_unpaid_work_yesno == 2

replace missed_unpaid_work_days = 0 if missed_unpaid_work_yesno == 0

///// EQ-5D-5L \\\\\ 1=no problems at all, ..., 5=extremely problematic

///// MHQoL \\\\\ 1=very well, ..., 4=very bad
drop PsychischwelzijnGeefopdeme
drop HE
drop HF

** create new variables needed in dataset
* generate BMI
replace height = 168 if ID == 130
gen BMI = weight/((height/100)^2)
order BMI, a(height)

** Drop observation ID == 94, since participant has outstanding number of gynaecologist visits for 6 months. Discusses
with professional. Since these are the only real visits she had, it is very unlikely that this is true --> extreme outlier.
drop if ID == 94

PRICES AND COSTS

** Jaarmutaties consumentenprijsindex(CPI): 2014-2015 (2015) = 0.6%; 2015-2016 (2016) = 0.3%; 2016-2017 (2017) =
1.4%; 2017-2018 (2018) = 1.7%; 2018-2019 (2019) = 2.6%; 2019-2020 (2020) = 1.3%
gen p_CPI_2014_2020 = 1.006*1.003*1.014*1.017*1.026*1.013
gen p_CPI_Belgian_study = 1.023 * 1.025 * 1.025 * 1.01 * 1.006*1.003*1.014*1.017*1.026*1.013
label var p_CPI_2014_2020 "CPI to convert prices from 2014 to 2020 prices"

** ALLE PRIJZEN ZIJN REFERENTIEPRIJZEN VOOR 2014.
** Referentieprijz (2014) "standaard consult huisarts"; Kostenhandleiding tabel 4.19
gen p2014_GP = 33
gen p2020_GP = p2014_GP * p_CPI_2014_2020
label var p2020_GP "Reference price for 1 GP visit (2020)"

** Referentieprijz (2014) "standaard consult huisarts"; Kostenhandleiding tabel 4.19
** Handleiding geeft aan zelfde prijs te gebruiken als huisarts aangezien niet beschikbaar voor POH (paragraaf 4.11).
gen p2014_POH = 33
gen p2020_POH = p2014_POH * p_CPI_2014_2020
label var p2020_POH "Reference price for 1 POH visit (2020)"

** Referentieprijz (2014) "maatschappelijk werk"; Kostenhandleiding paragraaf 4.15.2
gen p2014_social_worker = 65
gen p2020_social_worker = p2014_social_worker * p_CPI_2014_2020
label var p2020_social_worker "Reference price for 1 social worker visit (2021)"

** Referentieprijz (2014) "fysiotherapie"; Kostenhandleiding tabel 4.21
** All costs for physical therapist, Mensendieck, caesar therapist the same (€30-€34). Chose the physical therapist price
because stated as primary one in the question.
gen p2014_phys_therap = 33
gen p2020_phys_therap = p2014_phys_therap * p_CPI_2014_2020
label var p2020_phys_therap "Reference price for 1 physical therapist visit (2020)"

```

**** Referentieprij (2014) "" DOEN DIETIST**
**** No reference price available, searched on internet on different providers. What did they charge per appointment?**
**** All prices approximately between 63-73 euros per hour. Taken the average of 68. (www.independer.nl). Every appointment price of a dietician depends on time. Mostly, there are visits of 15, 30, 45 and 60 mins. So, taking the average of 30 mins leads to a price of €34 per visit.**
gen p2020_dietician = 34
label var p2020_dietician "Reference price for 1 dietician visit (2020)"

**** Referentieprij (2014) "" DOEN HOMEOPAAT**
**** No reference price available, searched on internet on different providers. What did they charge per appointment?**
**** Appointments vary (depending on being the first or not and between homeopath centers) between €55 and €110 (www.vereniginghomeopathie.nl). Average is €83.**
gen p2020_homeopath = 83
label var p2020_homeopath "Reference price for 1 homeopath visit (2020)"

**** Referentieprij (2014) "kostprijs per zitting eerstelijnspsycholoog"; Kostenhandleiding tabel 4.28**
gen p2014_psychologist = 64
gen p2020_psychologist = p2014_psychologist * p_CPI_2014_2020
label var p2020_psychologist "Reference price for 1 psychologist visit (2020)"

**** Referentieprij (2014) "" DOEN BEDRIJFSARTS**
**** Assumption: twice price of comparable service (GP visits). Twice because the service is longer than a standard GP appointment.**
gen p2020_occup_phys = 2*p2020_GP
label var p2020_occup_phys "Reference price for 1 occupational physician visits (2020)"

**** MEDICIJNEN - Hormonal therapy**

**** PHARMACIST COSTS**
**** €6.00 per prescribed medicine for maximum of 90 days.**
**** for 6 months:**
display = 365/2/90
**** 2.0277778**
gen p2014_pharmacist = 2.0277778*6
gen p2020_pharmacist = p2014_pharmacist * p_CPI_2014_2020
label var p2020_pharmacist "Reference price for pharmacist costs per prescribed medicine for 6 months (2020)"

gen p2014_first_pharm = 12
gen p2020_first_pharm = p2014_first_pharm * p_CPI_2014_2020
label var p2020_first_pharm "Reference price for pharmacist costs per prescribed medicine, first time (2020)"

**** Oral contraceptive = 1x per day (€0.01 per day) - farmacotherapeutischkompas.nl**
gen p2020_oral_contrac = 0.01*365/2
label var p2020_oral_contrac "Average reference price for oral contraception per person per 6 months"

**** IUD = 1x per 5 years (€0.07 per day) - farmacotherapeutischkompas.nl**
gen p2020_IUD = 0.07*365/2
label var p2020_IUD "Average reference price for intrauterine device per person per 6 months"

**** 3x per day 10 mg Provera (5 mg = 0.11); 2x per day 5 mg Orgametril (5 mg = €0.13) - farmacotherapeutischkompas.nl**
gen p2020_oral_progestat = ((0.11*6)+(0.13*2)/2)*365/2
label var p2020_oral_progestat "Average reference price for oral progestative per person per 6 months"

**** Lucrin = €3.64 per day; Zoladex = €3.77 per day - farmacotherapeutischkompas.nl**
gen p2020_GNRH_analog = ((3.64+3.77)/2)*365/2
label var p2020_GNRH_analog "Average reference price for GNRH analog per person per 6 months"

**** Referentieprij (2014) "spoedeisende hulp bezoek"; Kostenhandleiding paragraaf 4.4**
gen p2014_SEH_visit = 259
gen p2020_SEH_visit = p2014_SEH_visit * p_CPI_2014_2020
label var p2020_SEH_visit "Reference price for 1 SEH visits (2020)"

**** Referentieprij (2014) "ambulancerit"; Kostenhandleiding tabel 4.9**
gen p2014_ambulance = 515
gen p2020_ambulance = p2014_ambulance * p_CPI_2014_2020
label var p2020_ambulance "Reference price for going to hospital with an ambulance (2020)"

```

** Referentieprij (2014) "polikniekbezoek, gewogen gemiddelde"; Kostenhandleiding tabel 4.5
gen p2014_specialist_visits = 91
gen p2020_specialist_visits = p2014_specialist_visits * p_CPI_2014_2020
label var p2020_specialist_visits "Reference price for one specialist visit (2020)"

** BEHANDELINGEN
** Referentieprij (2014) "dagbehandeling in ziekenhuis"; Kostenhandleiding paragraaf 4.2
gen p2014_outpatient_treatment = 276
gen p2020_outpatient_treatment = p2014_outpatient_treatment * p_CPI_2014_2020
label var p2020_outpatient_treatment "Reference price for one outpatient treatment in hospital (2020)"

** Referentieprij (2014) "verpleegdag, gewogen gemiddelde"; Kostenhandleiding tabel 4.2
** inpatient care.
gen p2014_overnight_days = 476
gen p2020_overnight_days = (p2014_overnight_days * p_CPI_2014_2020)
label var p2020_overnight_days "Reference price for one inpatient care day (2020)"

** Referentieprij (2014) "Referentieprij huishoudelijk hulp"; Kostenhandleiding paragraaf 5.4
gen p2014_domestichelp = 14
gen p2020_domestichelp = (p2014_domestichelp * p_CPI_2014_2020)
label var p2020_domestichelp "Reference price for one hour of domestic help (2020)"

** Referentieprij (2014) "Referentieprij verzorgd worden"; Kostenhandleiding paragraaf 5.4
gen p2014_nursing = 14
gen p2020_nursing = (p2014_nursing * p_CPI_2014_2020)
label var p2020_nursing "Reference price for one hour of nursing (informal) (2020)"

** Referentieprij (2014) "Referentieprij praktische hulp"; Kostenhandleiding paragraaf 5.4
gen p2014_practicalhelp = 14
gen p2020_practicalhelp = (p2014_practicalhelp * p_CPI_2014_2020)
label var p2020_practicalhelp "Reference price for one hour of practical help (2020)"

** Referentieprij (2014) "Kosten per kilometer naar vervoermiddel"; Kostenhandleiding tabel 5.2
** Average price per km car and public transport
gen avg_price_km_2014 = 0.19
gen avg_price_km_2020 = (avg_price_km_2014 * p_CPI_2014_2020)
label var avg_price_km_2020 "Average price per km, extracted from price per km of car and public transport (2020)"

** Aanname: lopend vervoer brengt geen kosten met zich mee (tijd wat verwaarloosd kan worden volgens de handleiding)
gen p2020_transport_walking = 0
label var p2020_transport_walking "Average price per km, walking (2020) "

** Aanname: lopend vervoer brengt geen kosten met zich mee (tijd wat verwaarloosd kan worden volgens de handleiding)
gen p2020_transport_bicycle = 0
label var p2020_transport_bicycle "Average price per km, bicycle (2020)"

** Bij auto: maak gebruik van gemiddeld €3,- parkeerkosten.
gen p2014_transport_parkingcar = 3.00
gen p2020_transport_parkingcar = (p2014_transport_parkingcar * p_CPI_2014_2020)

label var p2020_transport_parkingcar "Price for parking (2020)"

** public transport - average price per km.

** Bij taxi: basistarief van €2.95 en €2.66 per km.
gen avgp_km_taxi_2014 = 2.66
gen avgp_km_taxi_2020 = (avgp_km_taxi_2014 * p_CPI_2014_2020)
label var avgp_km_taxi_2020 "Average price per km with taxi (2020)"

gen p2014_transport_tarifftax = 2.95
gen p2020_transport_tarifftax = (p2014_transport_tarifftax * p_CPI_2014_2020)
label var p2020_transport_tarifftax "Average price of start tariff with taxi (2020)"

** Gemiddelde afstanden andere locaties (km); Kostenhandleiding tabel 5.1
gen avg_distance_GP = 1.1
label var avg_distance_GP "Average distance from household to GP (km)"

```

```

gen avg_distance_phys_ther = 2.2
label var avg_distance_phys_ther "Average distance from household to physical therapist (km)"

** Average distance for all other facilities are assumed to be the same as the physical therapist
gen avg_distance_other = 2.2
label var avg_distance_other "Average distance from household to all other facilities (km) (assumed to be the same as
physical therapist)"

** Referentieprijis (2014) "Productiviteitskosten per uur betaald werkende"; Kostenhandleiding tabel 6.2
gen costs_2014_paidwork_women = 31.60
gen costs_2020_paidwork_women = costs_2014_paidwork_women * p_CPI_2014_2020
label var costs_2020_paidwork_women "Total costs of productivity loss for women in paid work (2020)"

** Referentieprijis (2014) "Productiviteitskosten per uur onbetaald werk"; Kostenhandleiding paragraaf 6.1.2.2.
gen costs_2014_unpaidwork = 14
gen costs_2020_unpaidwork = costs_2014_unpaidwork * p_CPI_2014_2020
label var costs_2020_unpaidwork "Total costs of productivity loss in unpaid work (2020)"

drop p2014_GP
drop p2014_POH
drop p2014_social_worker
drop p2014_phys_therap
drop p2014_psychologist
drop p2014_SEH_visit
drop p2014_ambulance
drop p2014_specialist_visits
drop p2014_outpatient_treatment
drop p2014_overnight_days
drop p2014_domestichelp
drop p2014_nursing
drop p2014_practicalhelp
drop avg_price_km_2014
drop avgp_km_taxi_2014
drop p2014_transport_parkingcar
drop p2014_transport_tarifftax
drop costs_2014_paidwork_women
drop costs_2014_unpaidwork

///// Utilities/QoL /////
** EQ-5D-5L: Application Dutch tariff (Matthijs Versteegh, 2015 (updated in 2018))
gen EQ5DTobit = 1.0
recast double EQ5DTobit

replace EQ5DTobit = EQ5DTobit -.0354544 if EQ_mobility==2
replace EQ5DTobit = EQ5DTobit -.0565962 if EQ_mobility==3
replace EQ5DTobit = EQ5DTobit -.166003 if EQ_mobility==4
replace EQ5DTobit = EQ5DTobit -.2032975 if EQ_mobility==5

replace EQ5DTobit = EQ5DTobit -.0381079 if EQ_selfcare==2
replace EQ5DTobit = EQ5DTobit -.0605347 if EQ_selfcare==3
replace EQ5DTobit = EQ5DTobit -.1677852 if EQ_selfcare==4
replace EQ5DTobit = EQ5DTobit -.1677852 if EQ_selfcare==5

replace EQ5DTobit = EQ5DTobit -.0391539 if EQ_daily_activ==2
replace EQ5DTobit = EQ5DTobit -.0867559 if EQ_daily_activ==3
replace EQ5DTobit = EQ5DTobit -.1924631 if EQ_daily_activ==4
replace EQ5DTobit = EQ5DTobit -.1924631 if EQ_daily_activ==5

replace EQ5DTobit = EQ5DTobit -.0658959 if EQ_pain==2
replace EQ5DTobit = EQ5DTobit -.0919691 if EQ_pain==3
replace EQ5DTobit = EQ5DTobit -.35993 if EQ_pain==4
replace EQ5DTobit = EQ5DTobit -.4152142 if EQ_pain==5

replace EQ5DTobit = EQ5DTobit -.069622 if EQ_fear==2
replace EQ5DTobit = EQ5DTobit -.1445222 if EQ_fear==3
replace EQ5DTobit = EQ5DTobit -.3563913 if EQ_fear==4
replace EQ5DTobit = EQ5DTobit -.4206361 if EQ_fear==5

```

```

gen any = 0
gen tot=(EQ_mobility+EQ_selfcare+EQ_daily_activ+EQ_pain+EQ_fear)
replace any=1 if tot>5

replace EQ5DTobit = EQ5DTobit -(1-.9530767) if any==1

gen missing=0
replace missing=1 if EQ_mobility==.
replace missing=1 if EQ_selfcare==.
replace missing=1 if EQ_daily_activ==.
replace missing=1 if EQ_pain==.
replace missing=1 if EQ_fear==.

replace EQ5DTobit = . if missing==1

drop any
drop tot
drop missing

label var EQ5DTobit "Total weighted EQ5D score"
label values EQ5DTobit EQ5DTobit

** MHQoL: Application
** Source: personal communication on Dutch MHQoL valueset

gen NL_MHQoL = 1.0
recast double NL_MHQoL

replace NL_MHQoL = NL_MHQoL -0.001539 if MHQoL_selfimage == 2
replace NL_MHQoL = NL_MHQoL -0.1347 if MHQoL_selfimage == 3
replace NL_MHQoL = NL_MHQoL -0.2063 if MHQoL_selfimage == 4

replace NL_MHQoL = NL_MHQoL -0.01301 if MHQoL_independence == 2
replace NL_MHQoL = NL_MHQoL -0.1166 if MHQoL_independence == 3
replace NL_MHQoL = NL_MHQoL -0.1809 if MHQoL_independence == 4

replace NL_MHQoL = NL_MHQoL -0.05137 if MHQoL_mood == 2
replace NL_MHQoL = NL_MHQoL -0.1739 if MHQoL_mood == 3
replace NL_MHQoL = NL_MHQoL -0.2993 if MHQoL_mood == 4

replace NL_MHQoL = NL_MHQoL -0.005415 if MHQoL_relationships == 2
replace NL_MHQoL = NL_MHQoL -0.1656 if MHQoL_relationships == 3
replace NL_MHQoL = NL_MHQoL -0.261 if MHQoL_relationships == 4

replace NL_MHQoL = NL_MHQoL -0.01248 if MHQoL_dailyactiv == 2
replace NL_MHQoL = NL_MHQoL -0.1343 if MHQoL_dailyactiv == 3
replace NL_MHQoL = NL_MHQoL -0.208 if MHQoL_dailyactiv == 4

replace NL_MHQoL = NL_MHQoL -0.05718 if MHQoL_physhealth == 2
replace NL_MHQoL = NL_MHQoL -0.2372 if MHQoL_physhealth == 3
replace NL_MHQoL = NL_MHQoL -0.3696 if MHQoL_physhealth == 4

replace NL_MHQoL = NL_MHQoL +0.01967 if MHQoL_future == 2
replace NL_MHQoL = NL_MHQoL -0.09776 if MHQoL_future == 3
replace NL_MHQoL = NL_MHQoL -0.162 if MHQoL_future == 4

** To correct for the one positive QALY-value (which can lead to an MHQoL score > 1).
replace NL_MHQoL = 1 if NL_MHQoL > 1

label var NL_MHQoL "Total weighted MHQoL score"
label values NL_MHQoL NL_MHQoL

** MHQoL: Application unweighted
** Source: MHQoL Scoring Manual
replace MHQoL_selfimage = 3 if MHQoL_selfimage == 1

```

```

replace MHQoL_selfimage = 2 if MHQoL_selfimage == 2
replace MHQoL_selfimage = 1 if MHQoL_selfimage == 3
replace MHQoL_selfimage = 0 if MHQoL_selfimage == 4

replace MHQoL_independence = 3 if MHQoL_independence == 1
replace MHQoL_independence = 2 if MHQoL_independence == 2
replace MHQoL_independence = 1 if MHQoL_independence == 3
replace MHQoL_independence = 0 if MHQoL_independence == 4

replace MHQoL_mood = 3 if MHQoL_mood == 1
replace MHQoL_mood = 2 if MHQoL_mood == 2
replace MHQoL_mood = 1 if MHQoL_mood == 3
replace MHQoL_mood = 0 if MHQoL_mood == 4

replace MHQoL_relationships = 3 if MHQoL_relationships == 1
replace MHQoL_relationships = 2 if MHQoL_relationships == 2
replace MHQoL_relationships = 1 if MHQoL_relationships == 3
replace MHQoL_relationships = 0 if MHQoL_relationships == 4

replace MHQoL_dailyactiv = 3 if MHQoL_dailyactiv == 1
replace MHQoL_dailyactiv = 2 if MHQoL_dailyactiv == 2
replace MHQoL_dailyactiv = 1 if MHQoL_dailyactiv == 3
replace MHQoL_dailyactiv = 0 if MHQoL_dailyactiv == 4

replace MHQoL_physhealth = 3 if MHQoL_physhealth == 1
replace MHQoL_physhealth = 2 if MHQoL_physhealth == 2
replace MHQoL_physhealth = 1 if MHQoL_physhealth == 3
replace MHQoL_physhealth = 0 if MHQoL_physhealth == 4

replace MHQoL_future = 3 if MHQoL_future == 1
replace MHQoL_future = 2 if MHQoL_future == 2
replace MHQoL_future = 1 if MHQoL_future == 3
replace MHQoL_future = 0 if MHQoL_future == 4

gen MHQoLScore=.
replace MHQoLScore = MHQoL_selfimage + MHQoL_independence + MHQoL_mood + MHQoL_relationships +
MHQoL_dailyactiv + MHQoL_physhealth + MHQoL_future

label var MHQoLScore "Unweighted total MHQoL score"
label values MHQoLScore MHQoLScore

** EHP-30 evaluation: source user manual of the Endometriosis Health Profile (EHP) - Oxford UK
(https://innovation.ox.ac.uk/outcome-measures/endometriosis-health-profile-ehp/)

** 1=never, ..., 5=always: the lower the score, the better.

** Scale pain
egen EHP30_points_pain = rowtotal(EHP30_skipsocials EHP30_nohousehold_act EHP30_pain_standing
EHP30_pain_sitting EHP30_pain_walking EHP30_no_sports EHP30_no_appetite EHP30_sleepless
EHP30_bed_due_pain EHP30_unable_dowhatwant EHP30_unbearable_pain)
label var EHP30_points_pain "Total points for EHP-30 score on the pain scale"

gen EHP30_score_pain = (EHP30_points_pain/55)*100
label var EHP30_score_pain "EHP-30 score on the pain scale (0-100)"

** Scale control and powerlessness
egen EHP30_points_control = rowtotal(EHP30_notfeelingwell EHP30_frustrated_no_improvements
EHP30_frustrated_no_control EHP30_unforgettable_symptoms EHP30_symptoms_control_life
EHP30_symptoms_hinder_life)
label var EHP30_points_control "Total points for EHP-30 score on the control and powerlessness scale"

gen EHP30_score_control = (EHP30_points_control/30)*100
label var EHP30_score_control "EHP-30 score on the control and powerlessness scale (0-100)"

** Scale emotional well-being

```

```

egen EHP30_points_wellbeing = rowtotal(EHP30_depressed EHP30_sad EHP30_miserable EHP30_moodswings
EHP30_moody EHP30_aggressive)
label var EHP30_points_wellbeing "Total points for EHP-30 score on the emotional well-being scale"

gen EHP30_score_wellbeing = (EHP30_points_wellbeing/30)*100
label var EHP30_score_wellbeing "EHP-30 score on the emotional well-being scale (0-100)"

** Scale social support
egen EHP30_points_support = rowtotal(EHP30_unable_tell_feelings EHP30_felt_misunderstood EHP30_felt_whiner
EHP30_felt_lonely)
label var EHP30_points_support "Total points for EHP-30 score on the social support scale"

gen EHP30_score_support = (EHP30_points_support/20)*100
label var EHP30_score_support "EHP-30 score on the social support scale (0-100)"

** Scale self-image
egen EHP30_points_selfimage = rowtotal(EHP30_not_desired_clothes EHP30_influenced_appearance
EHP30_no_confidence)
label var EHP30_points_selfimage "Total points for EHP-30 score on the self image scale"

gen EHP30_score_selfimage = (EHP30_points_selfimage/15)*100
label var EHP30_score_selfimage "EHP-30 score on the social self-image scale (0-100)"

CALCULATION COSTS
////// iMCQ \\\\\\\
** All costs are calculated for the year 2020 since this is the last available year for CPI and reference prices calculations.
** Total costs GP visits
gen TC_pp_GP = GP_visits * p2020_GP
label var TC_pp_GP "Total costs for GP visits per specific patient in past 6 months"

** Total costs POH visits
gen TC_pp_POH = POH_visits * p2020_POH
label var TC_pp_POH "Total costs for POH visits per specific patient in past 6 months"

** Total costs social worker visits
gen TC_pp_soc_worker = social_worker_visits * p2020_social_worker
label var TC_pp_soc_worker "Total costs for social worker visits per specific patient in past 6 months"

** Total costs physical therapist visits
gen TC_pp_phys_ther = phys_therap_visits * p2020_phys_therap
label var TC_pp_phys_ther "Total costs for physical therapist visits per specific patient in past 6 months"

** Total costs dietician visits - TO DO PRICE
gen TC_pp_dietician = dietician_visits * p2020_dietician
label var TC_pp_dietician "Total costs for dietician visits per specific patient in past 6 months"

** Total costs homeopath/acupuncturist visits - TO DO PRICE
gen TC_pp_homeopath = homeopath_visits * p2020_homeopath
label var TC_pp_homeopath "Total costs for homeopath/acupuncturist visits per specific patient in past 6 months"

** Total costs psychologist visits
gen TC_pp_psychologist = psychologist_visits * p2020_psychologist
label var TC_pp_psychologist "Total costs for psychologist visits per specific patient in past 6 months"

** Total costs occupational physician visits - TO DO PRICE
gen TC_pp_occ_phys = occupational_phys_visits * p2020_occup_phys
label var TC_pp_occ_phys "Total costs for occupational physician visits per specific patient in past 6 months"

** Total medicine costs - Hormonal therapy
gen TC_pp_oral_contrac = (p2020_oral_contrac * oral_contraceptive) + p2020_pharmacist if oral_contraceptive == 1
gen TC_pp_IUD = (p2020_IUD * hormonal_IUD) + p2020_first_pharm if hormonal_IUD == 1
gen TC_pp_oral_progestat = (p2020_oral_progestat * oral_progestativa) + p2020_pharmacist if oral_progestativa == 1
gen TC_pp_GNRH_analog = (p2020_GNRH_analog * GNRH_analoog) + p2020_pharmacist if GNRH_analoog == 1

** TC HORMONAL MEDICINES
egen TC_pp_medication = rowtotal(TC_pp_oral_contrac TC_pp_IUD TC_pp_oral_progestat TC_pp_GNRH_analog)

```

label var TC_pp_medication "Total medication costs, only pre-specified hormonal therapies used due to lack of data of others (negligible)"

** Total costs SEH visits

gen TC_pp_SEH = visits_SEH * p2020_SEH_visit

label var TC_pp_SEH "Total costs for SEH visits per specific patient in past 6 months"

** Total costs going to hospital with ambulance

gen TC_pp_ambulance = ambulance * p2020_ambulance

label var TC_pp_ambulance "Total costs for going to hospital with ambulance per patient in past 6 months"

** Total costs specialists visits

gen TC_pp_gynaecologist = gynaecologist * p2020_specialist_visits

label var TC_pp_gynaecologist "Total costs for gynaecologist visits per specific patient in past 6 months"

gen TC_pp_fert_special = fertility_specialist * p2020_specialist_visits

label var TC_pp_fert_special "Total costs for fertility specialist visits per specific patient in past 6 months"

gen TC_pp_specialist_visit = total_specialists_visits_pp * p2020_specialist_visits

label var TC_pp_specialist_visit "Total costs for specialists visits per specific patient in past 6 months"

gen TC_pp_other_spec_visits = total_other_spec_visits_pp * p2020_specialist_visits

label var TC_pp_other_spec_visits "Total costs for specialists visits (without gynaecologist and fertility specialist visits) per specific patient in past 6 months"

** Total costs outpatient care

gen TC_pp_outpatient = tot_outpatient_pp * p2020_outpatient_treatment

label var TC_pp_outpatient "Total costs for outpatient visits (summed visits)"

** Total costs inpatient care

gen TC_pp_inpatient = total_overnight_duration * p2020_overnight_days

label var TC_pp_inpatient "Total costs for inpatient care per specific patient in past 6 months"

** TOTAL PATIENT/FAMILY COSTS

** TRANSPORTATION COSTS

gen costs_2020_transport_walking = (p2020_transport_walking * (distance_hospital_km * 2)) if transport_walking == 1
label var costs_2020_transport_walking "Round trip travel costs to the hospital, walking (2020)"

gen costs_2020_transport_bicycle = (p2020_transport_bicycle * (distance_hospital_km * 2)) if transport_bicycle == 1
label var costs_2020_transport_bicycle "Round trip travel costs to the hospital, bicycle (2020)"

gen costs_2020_transport_car = ((avg_price_km_2020 * (distance_hospital_km * 2)) + p2020_transport_parkingcar) if transport_car == 1
label var costs_2020_transport_car "Round trip travel costs to the hospital by car, parking costs included (2020)"

gen costs_2020_transport_public = (avg_price_km_2020 * (distance_hospital_km * 2)) if transport_public == 1
label var costs_2020_transport_public "Round trip travel costs to the hospital by public transport (2020)"

gen costs_2020_transport_taxi = (avgp_km_taxi_2020 * (distance_hospital_km * 2)) + (2 * p2020_transport_tariff) if transport_taxi == 1
label var costs_2020_transport_taxi "Round trip travel costs to the hospital by taxi (2020)"

** Total hospital visits: SEH excluded since we don't know to what hospital and if it was with ambulance that time.

egen total_hospital_visits = rowtotal(total_specialists_visits_pp tot_outpatient_pp hospital_overnight_times visits_SEH)

label var total_hospital_visits "Total visits to hospital per patient in past 6 months"

egen TC_2020_transport_hospital = rowtotal(costs_2020_transport_walking costs_2020_transport_bicycle

costs_2020_transport_car costs_2020_transport_public costs_2020_transport_taxi)

label var TC_2020_transport_hospital "Total round trip travel costs to the hospital for 1 visit (summed over all ways of transport) (2020)"

gen total_ways_transport = .

replace total_ways_transport = 1 if ID == 130 | ID == 17 | ID == 140 | ID == 191 | ID == 216 | ID == 129 | ID == 203 | ID == 187 | ID == 299 | ID == 222 | ID == 44 | ID == 236 | ID == 5 | ID == 214 | ID == 243 | ID == 215 | ID == 174 | ID == 83 | ID == 8 | ID == 286 | ID == 20 | ID == 169 | ID == 66 | ID == 287 | ID == 296 | ID == 4 | ID == 205 | ID == 300 | ID == 12 | ID == 261 | ID == 53 | ID == 91 | ID == 151 | ID == 278 | ID == 59 | ID == 49 | ID == 36 | ID == 92 | ID == 253 | ID == 134 | ID == 302 | ID == 118 | ID == 166 | ID == 220 | ID == 212 | ID == 113 | ID == 259 | ID == 244 | ID == 149 | ID == 208 | ID == 161 | ID == 276 | ID == 254 | ID == 172 | ID == 304 | ID == 73 | ID == 260 | ID == 96 | ID == 146 | ID == 229 | ID == 173 | ID == 179 | ID == 31 | ID == 10 | ID == 29 | ID == 79 | ID == 239 | ID == 7 | ID == 126 | ID == 228 | ID == 111 | ID == 138 | ID == 238 | ID == 148 | ID == 71 | ID == 9 | ID == 135 | ID == 262 | ID == 241 | ID == 2 | ID

```

== 275 | ID == 190 | ID == 37 | ID == 120 | ID == 284 | ID == 70 | ID == 16 | ID == 178 | ID == 319 | ID == 122 | ID ==
221 | ID == 153 | ID == 183 | ID == 257 | ID == 250 | ID == 157 | ID == 230 | ID == 181 | ID == 100 | ID == 242 | ID ==
141 | ID == 67 | ID == 232 | ID == 206 | ID == 293 | ID == 43 | ID == 170 | ID == 248 | ID == 57 | ID == 280 | ID == 94 |
ID == 273 | ID == 81 | ID == 6 | ID == 303 | ID == 50 | ID == 85 | ID == 193 | ID == 117 | ID == 210
replace total_ways_transport = 2 if ID == 234 | ID == 27 | ID == 48 | ID == 305 | ID == 142 | ID == 121 | ID == 106 | ID
== 61 | ID == 14 | ID == 39 | ID == 245 | ID == 264 | ID == 32 | ID == 271 | ID == 315 | ID == 207
replace total_ways_transport = 3 if ID == 58 | ID == 168 | ID == 103
label var total_ways_transport "Total number of ways of transport filled in in the survey"

gen TC_2020_avg_trans_hosp = TC_2020_transport_hospital/total_ways_transport
label var TC_2020_avg_trans_hosp "Total average round trip costs to the hospital for 1 visit (divided by number of ways
of transport to get average) (2020)"

gen TC_ppavg_travelcosts_hospital = TC_2020_avg_trans_hosp * total_hospital_visits
label var TC_ppavg_travelcosts_hospital "Total average round trip travel costs per patient to hospital with all ways of
transportation in past 6 months"

** Since way of transport is not known to other facilities than hospital, price per km of car and public transport is taken
(Richtlijn voor het uitvoeren van economische evaluaties in gezondheidszorg)
gen TC_pp_travelcosts_GP = ((avg_distance_GP*2) * avg_price_km_2020 * (GP_visits+POH_visits))
label var TC_pp_travelcosts_GP "Total travel costs per patient to GP in past 6 months (taken average for ways of
transportation)"

gen TC_pp_travel_phys_ther = ((avg_distance_phys_ther*2) * avg_price_km_2020 * phys_therap_visits)
label var TC_pp_travel_phys_ther "Total travel costs per patient to physical therapist in past 6 months (taken average for
ways of transportation)"

gen TC_pp_travel_soc_worker = ((avg_distance_other*2) * avg_price_km_2020 * social_worker_visits)
label var TC_pp_travel_soc_worker "Total travel costs per patient to social worker in past 6 months (taken average for
ways of transportation)"

gen TC_pp_travel_dietician = ((avg_distance_other*2) * avg_price_km_2020 * dietician_visits)
label var TC_pp_travel_dietician "Total travel costs per patient to dietician in past 6 months (taken average for ways of
transportation)"

gen TC_pp_travel_homeopath = ((avg_distance_other*2) * avg_price_km_2020 * homeopath_visits)
label var TC_pp_travel_homeopath "Total travel costs per patient to homeopath/acupuncturist in past 6 months (taken
average for ways of transportation)"

gen TC_pp_travel_psychologist = ((avg_distance_other*2) * avg_price_km_2020 * psychologist_visits)
label var TC_pp_travel_psychologist "Total travel costs per patient to psychologist in past 6 months (taken average for
ways of transportation)"

gen TC_pp_travel_occ_phys = ((avg_distance_other*2) * avg_price_km_2020 * occupational_phys_visits)
label var TC_pp_travel_occ_phys "Total travel costs per patient to occupational physician in past 6 months (taken average
for ways of transportation)"

egen TOT_TRAVEL_COSTS = rowtotal(TC_ppavg_travelcosts_hospital TC_pp_travelcosts_GP
TC_pp_travel_phys_ther TC_pp_travel_soc_worker TC_pp_travel_dietician TC_pp_travel_homeopath
TC_pp_travel_psychologist TC_pp_travel_occ_phys)
label var TOT_TRAVEL_COSTS "Total travel costs per patient in past 6 months (summed over all facilities)"

** SUPPORTING ACTIVITIES IN HOUSEHOLD COSTS
gen TC_pp_domestichelp = total_domestichelp * p2020_domestichelp
label var TC_pp_domestichelp "Total costs for support in household activities, domestic help, for 6 months"

gen TC_pp_nursing = total_nursing * p2020_nursing
label var TC_pp_nursing "Total costs for support in household activities, nursing, for 6 months"

gen TC_pp_practicalhelp = total_practicalhelp * p2020_practicalhelp
label var TC_pp_practicalhelp "Total costs for support in household activities, practical help, for 6 months"

egen TOT_SUPP_HH_ACTIVITIES = rowtotal(TC_pp_domestichelp TC_pp_nursing TC_pp_practicalhelp)
label var TOT_SUPP_HH_ACTIVITIES "Total costs for support in household activities for 6 months"

```

////// iPCQ \\\\\\\

```

** PAID WORK - ABSENTEISM
** FRICTION COST METHOD
* filled vacancies, CBS vacancies for 2020 (*1000)
gen filled_vacancies_2020 = 1085
label var filled_vacancies_2020 "Number of filled vacancies in 2020"
* open vacancies, CBS vacancies for 2020 (*1000)
gen open_vacancies_2020 = 221.3
label var open_vacancies_2020 "Number of open vacancies in 2020"
* generate length vacancies
gen length_vacancies = 365 / (filled_vacancies_2020/open_vacancies_2020)
label var length_vacancies "Length of vacancies needed for friction cost period"
* calculate friction period in days by adding 4 weeks to the average length of vacancies.
gen friction_period_days = length_vacancies + 28
* Friction period in days was in 2019 111 days (Leona Hakkaart-van Roijen)
* below or above friction period
gen above_friction = .
replace above_friction = 1 if calendardays_long_absence > friction_period_days
replace above_friction = 0 if missing(calendardays_long_absence)
replace above_friction = 0 if calendardays_long_absence < friction_period_days
order above_friction, a(absenteism_long)
label var above_friction "Being absent from work more than 4 weeks and above the friction period"
label define above_friction 1 "Length of long absenteism above friction period" 0 "Length of long absenteism not above
friction period"

** SHORT ABSENTEISM
gen costs_short_absenteism = total_absenteism_hours*costs_2020_paidwork_women if absenteism_days > 0 |
absenteism_days != .
order costs_short_absenteism, a(total_absenteism_hours)
label var costs_short_absenteism "Costs of productivity loss with short absenteism over 4 weeks"
gen yearly_costs_short_absenteism = costs_short_absenteism*13
label var yearly_costs_short_absenteism "Costs annual productivity loss with short absenteism"

** LONG ABSENTEISM - SHORTER THAN FRICTION PERIOD
gen prod_loss_abs_SF = ((friction_period_days-calendardays_long_absence)/7)*paid_work_days*hours_per_workday if
above_friction==0 & absenteism_long==1
label var prod_loss_abs_SF "Productivity loss with long absenteism, but shorter than friction period"

gen costs_long_abs_SF = .
replace costs_long_abs_SF = (prod_loss_abs_SF*costs_2020_paidwork_women) if above_friction==0 &
absenteism_long==1
order costs_long_abs_SF, a(calendardays_long_absence)
label var costs_long_abs_SF "Costs of productivity loss with long absenteism, but shorter than friction period"

** LONG ABSENTEISM - LONGER THAN FRICTION PERIOD - productivity loss costs are 0
gen prod_loss_abs_LF = ((friction_period_days/7)*paid_work_days*hours_per_workday)
label var prod_loss_abs_LF "Productivity loss with long absenteism and longer than friction period due to endometriosiis"
gen costs_long_abs_LF = 0 if above_friction==1
label var costs_long_abs_LF "Costs of productivity loss with long absenteism and longer than friction period due to
endometriosiis"

** PAID WORK - PRESENTEEISM
gen prod_loss_presenteeism = presenteeism_days*(1-(score_productiv_workwhensymptoms/10))*hours_per_workday
label var prod_loss_presenteeism "Productivity loss during work because of pain due to endometriosiis (presenteeism)"
gen costs_presenteeism = prod_loss_presenteeism*costs_2020_paidwork_women
order costs_presenteeism, a(score_productiv_workwhensymptoms)
label var costs_presenteeism "Costs associated with productivity loss during work because of pain due to endometriosiis
(presenteeism)"
gen yearly_costs_presenteeism = costs_presenteeism*13
label var yearly_costs_presenteeism "Annual costs associated with productivity loss during work because of pain due to
endometriosiis (presenteeism)"

** UNPAID WORK
gen prod_loss_unpaidw = missed_unpaid_work_days*hours_replace_unpaidw_family
label var prod_loss_unpaidw "Productivity loss in unpaid work due to endometriosiis"
gen costs_loss_unpaidw = prod_loss_unpaidw*costs_2020_unpaidwork
order costs_loss_unpaidw, a(hours_replace_unpaidw_family)

```

```

label var costs_loss_unpaidw "Costs associated with loss in productivity in unpaidwork due to endometriosis"
gen yearly_costs_loss_unpaidw = costs_loss_unpaidw*13
label var yearly_costs_loss_unpaidw "Annual costs associated with loss in productivity in unpaidwork due to
endometriosis"

** TOTAL PRODUCTIVITY LOSS COSTS (PAID + UNPAID WORK)
egen YEAR_TOT_PROD_COSTS = rowtotal(yearly_costs_short_absenteism costs_long_abs_SF costs_long_abs_LF
yearly_costs_presenteeism yearly_costs_loss_unpaidw)
label var YEAR_TOT_PROD_COSTS "Total productivity loss costs including paid and unpaid work over a year"

** DIRECT HEALTH CARE COSTS
*** Yearly physician visits
gen yearly_GP = TC_pp_GP*2
label var yearly_GP "Yearly total GP costs"

gen yearly_POH = TC_pp_POH*2
label var yearly_POH "Yearly total POH costs"

gen yearly_GP_POH = yearly_POH + yearly_GP
label var yearly_GP_POH "Yearly total GP+POH costs"

gen yearly_gynaecologist_visits = TC_pp_gynaecologist*2
label var yearly_gynaecologist_visits "Yearly total gynaecologist costs"

gen yearly_fert_special = TC_pp_fert_special*2
label var yearly_fert_special "Yearly total fertility specialist costs"

gen yearly_specialist_visits = TC_pp_specialist_visit*2
label var yearly_specialist_visits "Yearly total specialist visits costs"

gen yearly_SEH = TC_pp_SEH*2
label var yearly_SEH "Yearly total SEH visits costs"

gen yearly_physician_visits = (yearly_GP + yearly_POH + yearly_specialist_visits + yearly_SEH)
label var yearly_physician_visits "Yearly total physician costs (GP, POH, specialist and SEH visits)"

gen yearly_other_spec_visits = (TC_pp_other_spec_visits*2)+yearly_SEH
label var yearly_other_spec_visits "Yearly total physician costs (specialist visits (without gynaecologist and fertility
specialist) and SEH visits)"

*** Yearly other visits
gen yearly_soc_worker = TC_pp_soc_worker*2
label var yearly_soc_worker "Yearly social worker visits costs"

gen yearly_phys_ther = TC_pp_phys_ther*2
label var yearly_phys_ther "Yearly physical therapist visits costs"

gen yearly_dietician = TC_pp_dietician*2
label var yearly_dietician "Yearly dietician visits costs"

gen yearly_homeopath = TC_pp_homeopath*2
label var yearly_homeopath "Yearly homeopath visits costs"

gen yearly_psychologist = TC_pp_psychologist*2
label var yearly_psychologist "Yearly psychologist visits costs"

gen yearly_occ_phys = TC_pp_occ_phys*2
label var yearly_occ_phys "Yearly occupational physician visits costs"

gen yearly_other_visits = (yearly_soc_worker + yearly_phys_ther + yearly_dietician + yearly_homeopath +
yearly_psychologist + yearly_occ_phys)
label var yearly_other_visits "Yearly total other visits costs (social worker, physical therapist, dietician, homeopath,
psychologist, occupational physician)"

*** Yearly treatments
gen yearly_treatments = (TC_pp_outpatient) * 2
label var yearly_treatments "Yearly total outpatient treatment costs"

```

```

gen yearly_hospitalization = (TC_pp_inpatient) * 2
label var yearly_hospitalization "Yearly total hospitalization costs (inpatient visits)"

*** Yearly medication
gen yearly_medication = (TC_pp_medication) * 2
label var yearly_medication "Yearly total medicine costs"

*** Yearly ambulance costs
gen yearly_ambulance = TC_pp_ambulance*2
label var yearly_ambulance "Yearly total ambulance transportation costs"

*** DIRECT HC COSTS
egen yearly_TOT_DIRECT_HCC = rowtotal(yearly_physician_visits yearly_other_visits yearly_treatments
yearly_hospitalization yearly_medication yearly_ambulance)
label var yearly_TOT_DIRECT_HCC "Yearly total direct health care costs"

** DIRECT NON-HC COSTS
gen yearly_travel_costs = TOT_TRAVEL_COSTS*2
label var yearly_travel_costs "Yearly total transportation costs (without ambulance)"
gen yearly_TOT_TRAVEL_COSTS = (yearly_travel_costs)
label var yearly_TOT_TRAVEL_COSTS "Yearly total transportation costs ( transportation (not ambulance))"
gen yearly_TOT_SUPP_HH_ACTIVITIES = TOT_SUPP_HH_ACTIVITIES * 2
label var yearly_TOT_SUPP_HH_ACTIVITIES "Yearly total costs for support in household activities"
egen yearly_TOT_DIRECT_NONHC = rowtotal(yearly_TOT_TRAVEL_COSTS
yearly_TOT_SUPP_HH_ACTIVITIES)
label var yearly_TOT_DIRECT_NONHC "Yearly direct non health care costs"

** DIRECT COSTS
egen yearly_DIRECT_COSTS = rowtotal(yearly_TOT_DIRECT_HCC yearly_TOT_DIRECT_NONHC)
label var yearly_DIRECT_COSTS "Yearly direct costs"

** INDIRECT COSTS
gen yearly_TOT_PROD_COSTS = YEAR_TOT_PROD_COSTS
label var yearly_TOT_PROD_COSTS "Yearly total productivity loss costs including paid and unpaid work"

** TOTAL COSTS
egen yearly_TOT_COSTS = rowtotal(yearly_DIRECT_COSTS yearly_TOT_PROD_COSTS)
label var yearly_TOT_COSTS "Yearly total costs related to endometriosis"

///// constructing total costs /////
*** Health care related costs
egen hc_related_costs_endo = rowtotal(TC_pp_GP TC_pp_POH TC_pp_soc_worker TC_pp_phys_ther TC_pp_dietician
TC_pp_homeopath TC_pp_psychologist TC_pp_occ_phys TC_pp_medication TC_pp_SEH TC_pp_ambulance
TC_pp_specialist_visit TC_pp_outpatient TC_pp_inpatient)
label var hc_related_costs_endo "Health care related costs of endometriosis (2020)"
gen yearly_HC_rel_costs = hc_related_costs_endo*2
label var yearly_HC_rel_costs "Yearly health care related costs of endometriosis (2020)"

** Total costs
egen yearly_total_costs = rowtotal(yearly_HC_rel_costs yearly_TOT_SUPP_HH_ACTIVITIES
yearly_TOT_TRAVEL_COSTS yearly_TOT_PROD_COSTS)
label var yearly_total_costs "Yearly total costs of endometriosis (2020)"

```

Appendix 3. Do-file assumption testing

** Conducted using Stata version 16.1

//// log_yearly_HC_rel_costs \\\

//// 1. Variable types

// a) all predictor variables must be quantitative or ctegrocial: CHECK.

// b) And outcome variable must be quantitative, continuous, and unbounded: CHECK.

//// 2. non-zero variance: predictors should have some variation in value (don't have variances of 0). CHECK

//// 3. no perfect multicollinearity: no perfect linear relationship between two or more predictors.

reg log_yearly_HC_rel_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

** VIF value: > 10 is wrong (0.1 as well) --> mean VIF is 3.49

vif

//// 4. Predictors are uncorrelated with external variables: CHECK

reg log_yearly_HC_rel_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

** hatsq should not be significant --> failed to reject assumption that model is specified correctly (of insignificant hatsq) = CHECK

linktest

** ovttest: we cannot reject H0 that model has no OVBs. = CHECK

ovttest

//// 5. Homoskedasticity: at each level of predictors, variance of residual terms should be constant. Homogeneity of variance: variances along line of best fit remain similar as you move along the line - plot residuals against predicted values

** plot of residuals versus fitted (predicted) values

** No strong evidence of heteroskedasticity = CHECK

rvfplot, yline(0)

** White's test: H0 = variance of residuals is homogeneous

** We cannot reject H0, so homogeneity (p=0.7777) = CHECK

estat imtest

//// 6. Independent errors: Durbin Watson test: not applicable for cross-sectional data.

//// 7. Normally distirbuted errors: normality, reduals (errors) should be normally distributed - histogram, normal P-P Plot, or normal Q-Q Plot of residuals = CHECHK

reg log_yearly_HC_rel_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

predict r, resid

kdensity r, normal

pnorm r

qnorm r

** door outliers in dataset - OK

** inter-quartile range (iqr): no severe outliers --> distribution seems fairly symmetric (only mild outliers which are common for samples of any size) = CHECK

iqr r

** Shapiro-Wilk W test: p = 0.22859: we cannot reject H0 that the distribution is normally distributed at a 5% significance level. = CHECK

swilk r

//// 8. independence = CHECK

```

//// 9. Linearity = CHECK
reg log_yearly_HC_rel_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

scatter r EQ5DTobit
scatter r NL_MHQoL
scatter r diagnostic_delay
scatter r grade_endo
scatter r BMI
scatter r age_cat
scatter r comorb_infertility

acprplot EQ5DTobit, lowess
acprplot NL_MHQoL, lowess
acprplot diagnostic_delay, lowess
acprplot grade_endo, lowess
acprplot BMI, lowess
acprplot age_cat, lowess
acprplot comorb_infertility, lowess

drop r

//// log_yearly_total_costs \\\
//// 1. Variable types
// a) all predictor variables must be quantitative or ctegrocial: CHECK.
// b) And outcome variable must be quantitative, continuous, and unbounded: CHECK.

//// 2. non-zero variance: predictors should have some variation in value (don't have variances of 0).

//// 3. no perfect multicollinearity: no perfect linear relationship between two or more predictors.
reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

** VIF value: > 10 is wrong (0.1 as well) --> mean VIF is 3.49
vif

//// 4. Predictors are uncorrelated with external variables
reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

** hatsq should not be significant --> have to reject assumption that model is specified correctly (significant hatsq) = NOT
CHECK
linktest

** ovtest: we have to reject H0 that model has no OVBs. = NOT CHECK
ovtest

//// 5. Homoskedasticity:at each level of predictors, variance of residual terms should be constant. Homogeneity of
variance: variances along line of best fit remain similar as you move along the line - plot residuals against predicted values
reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

** plot of residuals versus fitted (predicted) values
** No strong evidence of heteroskedasticity = CHECK
rvfplot, yline(0)

** White's test: H0 = variance of residuals is homogeneous
** We cannot reject H0, so homogeneity (p=0.9290) = CHECK
estat imtest

//// 6. Independent errors: Durbin Watson test: not applicable for cross-sectional data.

```

//// 7. Normally distributed errors: normality, residuals (errors) should be normally distributed - histogram, normal P-P Plot, or normal Q-Q Plot of residuals = CHECK
reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

predict r, resid
kdensity r, normal
pnorm r
qnorm r

** inter-quartile range (iqr): no severe outliers --> distribution seems fairly symmetric (only mild outliers which are common for samples of any size) = CHECK
iqr r

** Shapiro-Wilk W test: $p = 0.45451$: we cannot reject H_0 that the distribution is normally distributed at a 5% significance level. = CHECK
swilk r

//// 8. independence: assumed that all values of the outcome variable are independent. Each value of outcome variable comes from separate entity. = CHECK

//// 9. Linearity = CHECK

reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

predict r, resid

scatter r EQ5DTobit
scatter r NL_MHQoL
scatter r diagnostic_delay
scatter r grade_endo
scatter r BMI
scatter r age_cat
scatter r comorb_infertility

acprplot EQ5DTobit, lowess
acprplot NL_MHQoL, lowess
acprplot diagnostic_delay, lowess
acprplot grade_endo, lowess
acprplot BMI, lowess
acprplot age_cat, lowess
acprplot comorb_infertility, lowess

Appendix 4. Assumption testing; details

Table 14. Assumption testing: detail

Assumption	Explanation	Test	Outcome health care costs	Outcome total costs	Consequence
Non-zero variance	Predictors should not have variance of 0.	N/A	N/A	N/A	OK
No perfect multicollinearity	Predictors should have no distinct correlation with each other.	<i>VIF</i> : should be below 10. Value above 1 may bias results of the model [73].	<i>Mean VIF</i> : 3.49 May be a possible threat.	<i>Mean VIF</i> : 3.49 May be a possible threat.	Possible threats: <ul style="list-style-type: none"> • Standard error increases: limitation of trustworthiness. • Hard to assign importance to single predictors (hard to separate effects). • Limits explanation of unique variance of added variables: limits size of multiple correlation R. [59]
No omitted variables	Predictor variables should be uncorrelated with other, excluded, variables.	<i>Linktest</i> : H ₀ : model is specified correctly H ₁ : H ₀ is not true <i>Ovtest</i> : H ₀ : model has no OVB's H ₁ : H ₀ is not true	<i>Linktest</i> : P = 0.726 > 0.05 We cannot reject H ₀ that model is specified correctly. <i>Ovtest</i> : P = 0.238 > 0.05 We cannot reject H ₀ that model has no OVB's.	<i>Linktest</i> : P = 0.046 < 0.05 We have to reject H ₀ that model is specified correctly. <i>Ovtest</i> : P = 0.014 < 0.05 We have to reject H ₀ that model has no OVB's.	Test is OK for outcome health care costs. Possible threats for outcome total cost: results may be unreliable, biased and in turn may lead to wrong conclusions.
Homoskedasticity	All independent variables should have somewhat the same variance.	<i>Rvplot</i> : this is a plot of the residuals on the fitted predicted values. <i>White's test</i> : H ₀ : variance of residuals is homogeneous H ₁ : H ₀ is not true	<i>Rvplot</i> : no strong evidence of heteroskedasticity. <i>White's test</i> : P = 0.778 > 0.05 We cannot reject H ₀ that variance of residuals is homogeneous.	<i>Rvplot</i> : no strong evidence of heteroskedasticity. <i>White's test</i> : P = 0.929 > 0.05 We cannot reject H ₀ that variance of residuals is homogeneous.	Tests are OK. If tests were not OK. Predictor(s) would be not the most efficient one(s), but results are still unbiased.
Independent errors	Error terms of two different observations should be 0.	N/A	N/A	N/A	N/A

Normal distribution of errors	Errors (residuals) should be random and normally distributed. Zero variance assumed between observed data and model.	K-density histogram, normal Q-Q plot, normal P-P plot. <hr/> <i>Shapiro-Wilk W test</i> H ₀ : distribution is normally distributed H ₁ : H ₀ is not true <hr/> <i>Iqr test</i> : if there are severe outliers distribution does not seem symmetric.	K-density histogram, normal Q-Q plot and normal P-P plot are OK. <hr/> <i>Shapiro-Wilk W test</i> P = 0.229 > 0.05 We cannot reject H ₀ that distribution is normally distributed. <hr/> <i>Iqr test</i> : no severe outliers, only mild outliers which are common for samples of any size.	K-density histogram, normal Q-Q plot and normal P-P plot are OK. <hr/> <i>Shapiro-Wilk W test</i> P = 0.455 > 0.05 We cannot reject H ₀ that distribution is normally distributed. <hr/> <i>Iqr test</i> : no severe outliers, only mild outliers which are common for samples of any size.	Tests are OK. If tests were not OK, model would be inadequate and errors are not random.
Independence of outcomes	Outcomes are assumed to be independent.	N/A	N/A	N/A	N/A
Linear relationship predictors and outcomes	Relationship between predictors and outcomes should be linear.	<i>Acprplots</i> : graph of an augmented part plus the residual.	<i>Acprplots</i> : Smoothened lines are somewhat close to the line of the OLS predictions.	<i>Acprplots</i> : Smoothened lines are somewhat close to the line of the OLS predictions.	Test is OK. If test was not OK, range of values may increase and precision of estimates is decreased. Inference may be affected.

Appendix 5. Do-file inferential statistics and post hoc analyses

```
** Conducted using Stata version 16.1

/// Regression analysis \\\

* Log transformation health care related costs
gen log_yearly_HC_rel_costs = log(yearly_HC_rel_costs)
label var log_yearly_HC_rel_costs "log-transformed health care related year costs of endometriosis (2020)"

* Log transformation total costs
gen log_yearly_total_costs = log(yearly_total_costs)
label var log_yearly_total_costs "log-transformed total year costs of endometriosis (2020)"

* Simple regression analysis OLS
reg log_yearly_HC_rel_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility
reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

/// Sensitivity Analysis \\\

** GRADE_ENDO = REMOVED from analysis.
reg log_yearly_HC_rel_costs EQ5DTobit NL_MHQoL diagnostic_delay BMI i.age_cat i.comorb_infertility
reg log_yearly_total_costs EQ5DTobit NL_MHQoL diagnostic_delay BMI i.age_cat i.comorb_infertility

** MHQoL = REMOVED from analysis
reg log_yearly_HC_rel_costs EQ5DTobit diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility
reg log_yearly_total_costs EQ5DTobit diagnostic_delay i.grade_endo BMI i.age_cat i.comorb_infertility

/// Anova (Analysis of variance) test \\\

** TEST FOR DIFFERENT AGE CATEGORIES
graph box log_yearly_HC_rel_costs, over(age_cat)
** medians for group 1, 2 and 3 are higher than group 4 --> expect higher means as well

oneway log_yearly_HC_rel_costs age_cat
** P = 0.0055 < 0.05: there is a statistically significant difference between the means of log yearly health care related
costs for each of the age categories.
** Barlett's test for equal variance: you don't want significance. P = 0.397 > 0.05 --> we cannot reject H0 that variances
are equal (so insignificant test) = good

oneway log_yearly_HC_rel_costs age_cat, tabulate
oneway log_yearly_HC_rel_costs age_cat, bonferroni tabulate

graph box log_yearly_total_costs, over(age_cat)
** medians for group 1, 2 and 3 are higher than group 4 --> expect higher means as well

oneway log_yearly_total_costs age_cat
** P = 0.0127 < 0.05: there is a statistically significant difference between the means of log total health costs for each of
the age categories.
** Barlett's test for equal variance: you don't want significance. P = 0.092 > 0.05 --> we cannot reject H0 that variances
are equal (so insignificant test) = good

oneway log_yearly_total_costs age_cat, tabulate
oneway log_yearly_total_costs age_cat, bonferroni tabulate

** TEST FOR DIFFERENT GRADES OF ENDOMETRIOSIS
graph box log_yearly_HC_rel_costs, over(grade_endo)
** all medians approximately the same

oneway log_yearly_HC_rel_costs grade_endo
** P = 0.4693 > 0.05: there is a statistically insignificant difference between the means of log yearly health care related
costs for each of the grade endometriosis categories.
```

** Barlett's test for equal variance: you don't want significance. $P = 0.925 > 0.05$ --> we cannot reject H_0 that variances are equal (so insignificant test) = good

graph box log_yearly_total_costs, over(grade_endo)
** all medians approximately the same

oneway log_yearly_total_costs grade_endo

** $P = 0.8450 > 0.05$: there is a statistically insignificant difference between the means of log total health costs for each of the grade endometriosis categories.

** Barlett's test for equal variance: you don't want significance. $P = 0.339 > 0.05$ --> we cannot reject H_0 that variances are equal (so insignificant test) = good

Appendix 6. Topic list focus group sessions

Table 15. Topic list focus group sessions

Elements	Subjects	Prompts
Introduction	<ul style="list-style-type: none"> • Introducing myself and Leonie • Background and aims • Information on research (general, anonymity and safe, voluntary) • Verbal consent • Introducing participants 	N/A
Onset of endometriosis in life	Symptoms	<ul style="list-style-type: none"> • What impact did the first symptoms of endometriosis have on daily life? Can you explain this in detail? • What is your experience with the quest to answers and solutions to these symptoms? • What is your experience with healthcare professionals in this process?
	Diagnosis	<ul style="list-style-type: none"> • What is your experience with getting diagnosed? • Did you have any delay in diagnosis and how did this impact your mental well-being? • What is your experience with receiving information on endometriosis from healthcare professionals? • What is your experience with frustrations surrounding the delay in diagnosis or being misunderstood?
Life with endometriosis	Emotional impact	<ul style="list-style-type: none"> • What is the emotional impact of the diagnosis? • What impact did endometriosis have on your mental wellbeing in daily life?
	Impact on daily life	<ul style="list-style-type: none"> • What impact does endometriosis have on your social life? • What impact does endometriosis have on your studies? • What impact does endometriosis have on your work? • What financial impact does endometriosis have on your life? • What impact does endometriosis have on your lifestyle • What impact does endometriosis have on your intimate or sexual relationships and marriage?
	Support and understanding	<ul style="list-style-type: none"> • What is your experience with health care professionals and their support? • What is your experience with medical treatments (e.g. in hospitals)? Did this reduce your symptoms? • What is your experience with treatments concerning lifestyle (e.g. diet)? Did this reduce your symptoms? • What is your experience with alternative treatments (e.g. acupuncturist)? Did this reduce your symptoms? • Did you always feel understood/heard by health care professionals? • Did you always feel understood/heard by relatives? • What is your experience with knowledge of endometriosis by partners, family, friends and the government ? • What is your experience with knowledge and awareness of endometriosis by health care professionals?
New reality of living with endometriosis	New reality of living with endometriosis	<ul style="list-style-type: none"> • What does a typical week with endometriosis look like? • How does this typical week change when your period is about to take place?

		<ul style="list-style-type: none"> • What is your experience with modifications that should be made to deal with endometriosis? • Is there anything you (will have to) miss because of endometriosis?
Future with endometriosis	Future with endometriosis	<ul style="list-style-type: none"> • What is your opinion on your future with endometriosis? • What impact does endometriosis have on becoming a mother and possible infertility?
Motivation participating in session	Motivation participating in session	<ul style="list-style-type: none"> • What was your motivation to participate in these focus group sessions?
Conclusion	Concluding questions	<ul style="list-style-type: none"> • Is there anything you still want to say? • Is there anything you have missed during this session? • Do you have any questions for me/us? • Are you interested in receiving the summary of the focus group sessions and/or the article?
	Expressing gratitude for participating in the session	N/A