Ervaringen van vrouwen met (onverklaarde) gezondheidsklachten

Een quick scan

Sandra van Dulmen, Nivel
Merlijn van Breugel, Ditto Care
Peter Spreeuwenberg, Nivel
Mirjam Kaijer, Voices for Women
Roman Schut, Voices for Women

Het Nivel levert kennis om de gezondheidszorg in Nederland beter te maken. Dat doen we met hoogwaardig, betrouwbaar en onafhankelijk wetenschappelijk onderzoek naar thema's met een groot maatschappelijk belang. 'Kennis voor betere zorg' is onze missie. Met onze kennis dragen we bij aan het continu verbeteren en vernieuwen van de gezondheidszorg. We vinden het belangrijk dat mensen in staat zijn om deel te nemen aan de samenleving. Ons onderzoek draait uiteindelijk om de vraag hoe we de zorg voor de patiënt kunnen verbeteren. Alle onderzoeken publiceert het Nivel openbaar, dat is statutair vastgelegd.

Ditto Care bouwt een patiëntenoplossing in de zorg: een app die medische gesprekken opneemt en omzet in begrijpelijke samenvattingen, die je eenvoudig deelt met dierbaren. Daarachter zit een uitgebreide AI-motor die spraaktechnologie en taalmodellen combineert, met sterke nadruk op Responsible AI en het automatisch toetsen van de kwaliteit van elke samenvatting. Ditto is een Rotterdamse start-up uit 2024, won de nationale zorginnovatieprijs in 2026 en werd in de eerste 9 maanden na lancering al ruim 100.000 keer gedownload.

Stichting Voices for Women zet zich in voor een vrouwsensitieve gezondheidszorg waarin de man-vrouwverschillen worden erkend en meegenomen in diagnostiek, behandeling, onderzoek, onderwijs, beleid en politieke besluitvorming. De stichting geeft vrouwen met onverklaarde of onvoldoende herkende gezondheidsklachten een stem en zet hun ervaringen om in kennis, bewustwording en maatschappelijke verandering. Door vrouwengezondheid te monitoren en ervaringen van vrouwen te verzamelen en analyseren, draagt de stichting bij aan het zichtbaar maken van knelpunten en aan de verbetering van zorg, onderzoek en beleid, met oog voor de diversiteit onder vrouwen en de invloed van sociale, culturele en etnische factoren op gezondheid.

Stichting Voices for Women heeft van ZonMw financiering ontvangen voor het uitvoeren van een gecombineerd kwalitatief en kwantitatief onderzoek naar de ervaringen van vrouwen die via het meldpunt van de Stichting op www.voicesforwomen.nl binnen zijn gekomen. Nivel heeft deze opdracht aangenomen en samen met Ditto Care de analyses uitgevoerd.

Augustus 2026

030 272 97 00
nivel@nivel.nl
www.nivel.nl

© 2026 Nivel, Postbus 1568, 3500 BN UTRECHT

Gegevens uit deze uitgave mogen worden overgenomen onder vermelding van Nivel en de naam van de publicatie. Ook het gebruik van cijfers en/of tekst als toelichting of ondersteuning in artikelen, boeken en scripties is toegestaan, mits de bron duidelijk wordt vermeld.

Inhoud

1 Samenvatting 4

2 Inleiding 6
2.1 Doel 7
2.2 Onderzoeksvragen 7

3 Methode 8
3.1 Kwantitatieve analyse 8
3.2 Kwalitatieve analyse 9

4 Resultaten 13
4.1 Resultaten van de kwantitatieve analyse 13
4.2 Resultaten van de kwalitatieve analyse 18
4.3 Beknopte casusbeschrijvingen 28

5 Samenvatting en discussie 29
5.1 Samenvatting 29
5.2 Discussie 29

6 Literatuur 33

Bijlage A Registry met vragen per type analyse 35

Bijlage B Leren over LLM-beoordeling in thematische analyse 38

1 Samenvatting

Stichting Voices for Women (VfW) verzamelt sinds 2021 ervaringen van vrouwen met (onverklaarde) gezondheidsklachten. Deze ervaringen kunnen zowel door het invullen van een enquête als door middel van ervaringsverhalen gerapporteerd worden. Omdat deze ervaringen spontaan via de website van de Stichting worden aangeleverd, betreft het hier een zelf-geselecteerde - niet representatieve - groep vrouwen. De omvang en variëteit aan ervaringen die zijn gedeeld bieden niettemin een waardevolle getuigenis van door vrouwen ervaren knelpunten en hiaten in de Nederlandse gezondheidszorg. De verzamelde ervaringen zijn in opdracht van ZonMw zowel kwantitatief als kwalitatief geanalyseerd ten behoeve van de Werkagenda Vrouwengezondheid van VWS.

Voor de kwantitatieve analyse is gebruik gemaakt van twee van de drie beschikbare databestanden met resultaten van verschillende versies van een enquête waarmee vrouwen hun ervaringen met de gezondheidszorg konden delen via de website van Voices for Women. Deze twee versies (*Enquête-2* en *Enquête-3)* die door 368 respectievelijk 1269 vrouwen (totaal 1637) waren ingevuld, bevatten de meest gelijksoortige vragen; de antwoorden op de derde en oudste versie van de enquête (*Enquête-1)* met grotendeels afwijkend geformuleerde en merendeels open vragen die door 2268 vrouwen was ingevuld, is gebruikt voor een validatie van de resultaten van de enquête op vergelijkbare onderdelen.

De gemiddelde leeftijd van de 1637 respondenten was 50,5 jaar (range 15-86 jaar), de gemiddelde leeftijd waarop hun gezondheidsklachten begonnen was 33,5 jaar (range 0-76 jaar). 60% van de respondenten liep al 10 jaar of langer met de gezondheidsklachten rond. De top 5 van de meest gerapporteerde lichamelijke gezondheidsklachten omvat: vermoeidheid (ervaren door 77% van de vrouwen die de enquête heeft ingevuld); pijnlijke gewrichten (57%); spijsverteringsklachten (44%); onverklaarde pijn (43%); en, spierpijn (43%). De top 5 van meest gerapporteerde psychische gezondheidsklachten: slaapproblemen (58%); concentratieverlies (52%); stemmingswisselingen (51%); somberheid (49%); en, stress (49%). Burn-out/stress, de overgang en depressie waren de meest gestelde diagnosen bij respectievelijk 34%, 25% en 24% van de vrouwen die de enquête hadden ingevuld. Vrouwen zijn het meest vaak verwezen naar de gynaecoloog (42%), psycholoog (39%) en internist (29%). Kijkend naar de verschillen in klachten en diagnosen tussen de grootste leeftijdscategorieën (25-49 jaar en 50-65 jaar) laat zien dat er in de top 5 van psychische klachten weinig verschil is; de top 5 van lichamelijke klachten en diagnosen verschilt enigszins. De meeste vrouwen (75%) voelden zich - ongeacht hun leeftijd - bij de huisarts en de specialist niet gehoord. Nagenoeg alle respondenten gaven aan dat de klachten ernstige maatschappelijke gevolgen had. Het beeld dat uit de enquêtes naar voren kwam bleek vergelijkbaar met de antwoorden op de poll.

Voor de kwalitatieve analyse is een corpus van 1.285 ervaringsverhalen op twee manieren met behulp van AI geanalyseerd: deductief aan de hand van de door VfW opgestelde onderzoeksvragen c.q. thema's (de 'vragenregistry'; zie Bijlage A) en inductief via taalmodel-ondersteunde thematische codering volgens het vocabulaire van Braun en Clarke (2022). Ter validatie van de gegenereerde thema's zijn 20 ervaringsverhalen handmatig gecheckt.

Uit deze kwalitatieve analyses kwamen drie observaties consistent naar voren:

1. Het gebrek aan erkenning dat vrouwen ervaren is vrijwel universeel; in 99% van de verhalen worden de belastende aspecten van het zorg- en ziekteproces beschreven, en die draaien in meerderheid om het niet serieus genomen worden.
2. Psychologisering van lichamelijke klachten door de zorgverlener vormt de rode draad die in alle leeftijdsgroepen terugkomt.
3. Het inductieve spoor brengt patronen aan het licht waar de registry niet expliciet naar vroeg, in het bijzonder het verlies van werk, identiteit en levenskwaliteit als maatschappelijk-economisch gevolg van de gezondheidsklachten en de zoektocht naar een diagnose.

De resultaten van dit onderzoek leveren primair aanknopingspunten op voor een van de drie actielijnen van de VWS Werkagenda vrouwengezondheid: Het herkennen en erkennen van klachten. Slechts één op de vier vrouwen die hun ervaringen hebben gedeeld door middel van het invullen van de enquête geeft in dit onderzoek immers aan zich gehoord te hebben gevoeld bij de zorgverlener, het merendeel heeft negatieve ervaringen en voelt zich weggezet en hun klachten gebagatelliseerd. Deze uitkomsten en die van meer uitgebreid vervolgonderzoek kunnen gebruikt worden om zorgverleners meer bewust te maken van deze ervaringen.

2 Inleiding

In februari 2026 is een convenant ondertekend als opstap naar een Nationale Strategie Vrouwengezondheid (2025-2030). Daarmee zet het ministerie van VWS zich samen met vele andere partijen in op structurele verbeteringen in gezondheid en zorg voor vrouwen en meisjes. Het doel is dat alle vrouwen en meisjes in Nederland tijdig passende zorg en ondersteuning ontvangen en de gezondheidsverschillen tussen vrouwen en mannen kleiner worden. Op die manier kunnen vrouwen langer in goede gezondheid leven en volwaardiger deelnemen aan de samenleving. Bij de uitwerking van de strategie worden de volgende uitgangspunten gehanteerd:

- Vrouwengezondheid gaat over lichamelijke, mentale én sociale gezondheid.
- Elke levensfase kent specifieke uitdagingen. Vrouwengezondheid vraagt daarom om een levensloopbenadering waarin elke levensfase passende aandacht krijgt.
- Dé vrouw bestaat niet, vrouwen zijn onderling verschillend. Bij vrouwengezondheid worden achtergrond, leefomstandigheden en gezondheidsuitkomsten nadrukkelijk meegenomen.

De Nationale Strategie Vrouwengezondheid zal zich specifiek richten op: meer kennis en innovatie van vrouwsensitieve zorg; tijdige zorg en ondersteuning en betere informatie op maat gedurende de levensloop; en, op bewustwording en het bundelen van krachten.

Meer bewustwording is onder meer van belang omdat onderzoek uitwijst dat, ondanks het bestaan van klinische richtlijnen, er verschillen zijn in het voorschrijf- en verwijsgedrag van mannelijke en vrouwelijke zorgverleners (Eggermont et al, 2018; 2022) waarbij het bovendien uitmaakt of de patiënt een man of een vrouw is (Eggermont et al, 2018; 2022; Noordman & van Dulmen, 2016). Dit wordt bevestigd in onderzoek van Ballering et al (2021) dat laat zien dat vrouwen in vergelijking met mannen minder vaak fysiek onderzocht worden en minder beeldvormend onderzoek en verwijzingen krijgen, maar wel meer laboratoriumonderzoek. Vrouwen krijgen bovendien voor hun fysieke symptomen minder vaak een ziektediagnose wat blijkt samen te hangen met het geringere aantal diagnostische interventies dat zij ondergaan. Ook wanneer wordt ingezoomd op specifieke klachten zoals rugpijn, vermoeidheid, klachten aan arm en/of been en tintelende ledematen worden bij vrouwen na diagnostisch onderzoek minder ziektediagnoses gesteld dan bij mannen (Ballering et al, 2022). Niettemin zijn er veel aandoeningen die vaker voorkomen bij vrouwen dan bij mannen en bovendien anders worden ervaren, zoals schildklieraandoeningen (Calissendorff et al, 2023).

Het langer onverklaard blijven van gezondheidsklachten bij vrouwen kan niet uitsluitend worden verklaard door verschillen in de manier waarop klachten worden ervaren of geuit. Een belangrijke onderliggende factor is dat het mannenlichaam jarenlang de medische norm is geweest in onderzoek, onderwijs en zorg. Hierdoor zijn ziektebeelden, symptomen en diagnostische criteria vooral ontwikkeld op basis van mannelijke populaties, terwijl vrouwen vaak andere symptomen, ziektepatronen en ziekteverloop laten zien. Dit kan ertoe leiden dat klachten van vrouwen minder snel worden herkend, verklaard of adequaat behandeld.

Onderzoek naar langdurig onverklaarde gezondheidsklachten laat zien dat artsen patiënten, ongeacht geslacht, ruimte geven om hun klachten te bespreken, maar dat de ideeën, zorgen en ervaringen van patiënten niet altijd worden meegenomen in de interpretatie en verklaring van die klachten (Olde Hartman et al, 2013). Patiënten ervaren hierdoor regelmatig een mismatch tussen hun eigen beleving en de medische duiding van hun klachten (Houwen et al, 2017). Daarbij spelen niet alleen verschillen in perspectief tussen patiënt en zorgverlener een rol, maar ook de

beperkingen van een zorgsysteem dat historisch onvoldoende rekening heeft gehouden met sekse en gendersensitieve verschillen in gezondheid.

Hoe vrouwen met onverklaarde gezondheidsklachten de gezondheidszorg ervaren en welke patronen, knelpunten en emoties daarin naar voren komen is nog weinig onderzocht. Inzicht in deze aspecten vormt echter belangrijke input voor de strategische aanpak van vrouwengezondheid. Stichting Voices for Women (VfW) zet zich in voor een vrouwsensitieve gezondheidszorg waarin verschillen tussen vrouwen en mannen worden erkend, onderzocht en meegenomen in diagnostiek, behandeling, onderzoek, onderwijs en beleid. Voices for Women geeft vrouwen met onverklaarde of onvoldoende herkende gezondheidsklachten - ongeacht hun achtergrond of herkomst - een stem en zet hun ervaringen om in kennis, bewustwording en maatschappelijke verandering. Via het meldpunt op www.voicesforwomen.nl ontvangen ze veel persoonlijke ervaringsverhalen van vrouwen met klachten, zowel in de vorm van op zichzelf staande verhalen als middels antwoorden op vragen als onderdeel van een enquête. Onderzoek naar deze ervaringen kan meer licht werpen op mogelijke oorzaken en patronen die ten grondslag liggen aan niet tijdige herkenning van onderliggende problematiek.

2.1 Doel

Dit onderzoek heeft als doel om trends en patronen zichtbaar te maken in de ervaringen met de gezondheidszorg van vrouwen met langdurige (onverklaarde) gezondheidsklachten. Daarbij wordt specifiek gekeken naar de drie domeinen fysiek, psychisch en sociaal en de rol van leeftijd. De resultaten dienen als input voor de Werkagenda Vrouwengezondheid van VWS. Vanwege de korte looptijd van het project gaat het primair om een quick scan.

2.2 Onderzoeksvragen

VfW heeft in samenspraak met ZonMw de volgende onderzoeksvragen opgesteld als richtsnoer voor de analyse van de data.

1. **Thematische analyse van ervaringen**
   Welke onderwerpen en patronen komen het meest voor in de ervaringsverhalen?
2. **Gezondheid en impact**
   Hoe uiten ervaringen zich binnen de drie domeinen: fysieke, mentale en sociale gezondheid?
3. **Sentimenten en beleving**
   Welke emoties en ervaringen domineren? Hoe verschillen deze tussen subgroepen en levensfasen?
4. **Knelpunten in het zorgproces**
   Waar lopen mensen tegenaan in de zorg? Zijn daar terugkerende patronen in te herkennen?
5. **Kwantitatieve hoofdlijnen**
   Wat zijn de belangrijkste verdelingen en trends in de data?
6. **Duiding en vertaling**
   Hoe verhouden de uitkomsten zich tot de actielijnen van de VWS Werkagenda Vrouwengezondheid?

3 Methode

In dit onderzoek is een combinatie van kwantitatieve en kwalitatieve analysemethoden gehanteerd. In Bijlage A zijn de te beantwoorden onderzoeksvragen nader uitgewerkt voor de verschillende typen analyse. Resultaten van de online enquêtes zijn kwantitatief geanalyseerd door het Nivel. De antwoorden op de open vragen uit deze enquête en de opgetekende ervaringsverhalen zijn met behulp van AI op een kwalitatieve manier thematisch geanalyseerd door Ditto Care. Daarmee kon een eerste duiding gegeven worden aan de meer kwantitatieve resultaten. Er werd verzocht om een gerichte quick scan op hoofdlijnen. De nadruk zou daarbij moeten liggen op het identificeren van patronen, knelpunten en behoeften. Daarnaast werd verzocht om vijf typerende casussen te presenteren. Vrouwen melden zich bij VfW op eigen initiatief, doorgaans omdat zij ervaren dat hun klachten in de reguliere zorg onvoldoende worden herkend. De steekproef is daarmee per definitie geen afspiegeling van de Nederlandse vrouwelijke populatie, maar van vrouwen die lang zoekende zijn (geweest) naar een diagnose en de urgentie voelden om hun verhaal vast te leggen of de enquête in te vullen. Deze zelfselectie is geen tekortkoming van de analyse maar een definiërende eigenschap van het corpus van ervaringsverhalen.

3.1 Kwantitatieve analyse

Voor de kwantitatieve analyse is gebruik gemaakt van gegevens afkomstig van twee van de drie beschikbare versies van de enquête (‘*Enquête-2’* ingevuld tussen mei 2024 en februari 2025 en ‘*Enquête-3’* ingevuld tussen april 2025 en mei 2026) waarmee vrouwen hun ervaringen met de gezondheidszorg konden delen. Deze versies, die door 363 respectievelijk 1274 vrouwen zijn ingevuld, bevatten de meest gelijksoortige vragen; antwoorden op de derde versie van de enquête (‘*Enquête-1’*) met afwijkend geformuleerde vragen die door 2268 vrouwen is ingevuld vanaf oktober 2021 tot 2023 (maand onbekend), is gebruikt voor validatie op onderdelen. De stichting heeft nog aanvullende data beschikbaar voor toekomstige analyses.

3.1.1 Data-analyse

Voor het beantwoorden van de onderzoeksvragen is op basis van de data uit *Enquête-2* en *Enquête-3* de variabele ‘leeftijd’ onderverdeeld in de vier leeftijdsgroepen: ≤24 jaar; 25 t/m 49 jaar; 50 t/m 65 jaar; en ≥66 jaar en de variabele ‘duur van de gezondheidsklachten’ in twee groepen: 0 t/m 9 jaar; en ≥10 jaar. Gezien de grote verschillen in de grootte van de leeftijdsgroepen (de hoogste en de laagste groep bevatten relatief weinig respondenten) zijn de analyses gericht op verschillen in leeftijd alleen gedaan voor leeftijdsgroep 2 en 3 en beide duurgroepen (tabel 1).

Tabel 1 *Uitsplitsing van data uit Enquête-2 en Enquête-3 naar leeftijd en duur van de klachten*

	Duur 0 t/m 9 jaar	**Duur ≥10 jaar**	**Totalen**
Leeftijd ≤24 jaar	41	9	50
Leeftijd 25 t/m 49 jaar	297	314	611
Leeftijd 50 t/m 65 jaar	283	514	797
Leeftijd ≥66 jaar	32	103	135
Totalen	653	940	1593[1]

[1] Doordat niet iedereen haar leeftijd heeft ingevuld wijkt dit aantal af van het aantal ingevulde enquêtes.

Gezien de verschillen in de formulering en/of antwoordmogelijkheden van de vragen in de verschillende vragenlijsten moesten een aantal hercoderingen in de variabelen worden toegepast om toch een vergelijking tussen vragenlijsten mogelijk te maken. Zo is niet alleen, zoals hierboven beschreven, de variabele leeftijd gehercodeerd, maar zijn ook de antwoordmogelijkheden “Ik voel(de) mij (bijna) altijd gehoord” en “Ik voel(de) mij (bijna) helemaal niet gehoord” op de vraag “In hoeverre voel(de) je je gehoord bij de huisarts/specialist?” uit *Enquête-3* samengenomen tot ‘ja’ respectievelijk ‘nee’, omdat in *Enquête-2* en *Enquête-1* deze vraag als ”Voel(de) je je gehoord bij de huisarts?” was geformuleerd met als antwoordmogelijkheden ‘ja’ en ‘nee’.

De resultaten zijn primair met beschrijvende statistiek geanalyseerd. Multivariate analyses waren als onderdeel van deze quick scan niet mogelijk als gevolg van de vele open vragen dan wel antwoordmogelijkheden op een vraag.

3.2 Kwalitatieve analyse

Voor de kwalitatieve analyse is gebruik gemaakt van 1.285 sinds 2021 verzamelde en door Marktonderzoeksbureau EM Onderzoek geanonimiseerde ervaringsverhalen (het ‘corpus') uit zes bronnen: vier Word-documenten met verhalen (totaal n=900), twee xlsx-formulieren met de meldpunt-vragenlijsten *Deel je verhaal* en *Vertel je verhaal* (totaal n=385) en inzendingen via de website (n=38). Zie tabel 2 voor de verdeling. Het totale tekstvolume bedroeg circa 344.000 woorden, verdeeld over ongeveer 3.500 ingevulde open-tekstvelden. Gestructureerde respondenten (n=565) leverden afzonderlijke antwoorden op gestandaardiseerde subvragen; vrije narratieven (n=720) bevatten één doorlopend tekstveld per persoon. De analyse gebruikte beide formats zonder voorbewerking.

Tabel 2 *Verdeling van het corpus over bronbestanden inclusief aantal woorden*

Bron	n	Woorden	Mediaan w/verhaal
Ervaringsverhalen-1 (Map met verhalen)	682	175.719	220
Ervaringsverhalen-2 (Deel je verhaal)	316	70.967	171
Ervaringsverhalen-3 (Verhalen)	138	38.178	227
Ervaringsverhalen-4 (Vertel je verhaal)	69	16.513	173
Ervaringsverhalen-5 (Verhalen)	42	14.695	269
Ervaringsverhalen-6 (Website VfW)	38	27.885	652
Totaal	**1.285**	**343.957**	

Verwerking vond plaats in Ditto’s beveiligde cloudomgeving binnen de EER (AES-256 *at rest*, TLS 1.3 *in transit*); data werden niet gebruikt voor modeltraining en aan het einde van het project verwijderd. Een DSA is beschikbaar. Het taalmodel dat de analyse uitvoert (Google Gemini 3.5 Flash) draaide op een EU-geadresseerd Vertex AI-eindpunt.

3.2.1 Gebruik van een taalmodel

Een taalmodel als Google Gemini 3.5 Flash is een statistisch model dat natuurlijke tekst kan lezen en, op basis van instructies, gestructureerde tekst terug kan geven. Ditto voedde het model per aanroep met één onderzoeksvraag en/of één verhaal samen met een *prompt*, een gestructureerde instructie in tekst die beschrijft welke taak het uitvoert, welk antwoordformaat werd verwacht (JSON), en welke regels gelden. Voor deze analyse was de volgende regel het strengst; het model mocht uitsluitend werken op basis van wat in het verhaal staat, zonder externe kennis of aannames, en

moest elke claim onderbouwen met een letterlijk citaat. De pijplijn herhaalde dit proces één verhaal tegelijk en bewaarde elke aanroep, met modelversie en promptversie, in een reproduceerbaar logbestand.

3.2.2 Plan van aanpak

1. Deductief scoren

Per (verhaal × onderzoeksvraag) deed de pijplijn één LLM-aanroep. Het model las de vraag en het verhaal, identificeerde relevante passages, citeerde ze letterlijk en gaf een gestructureerde score terug: of de vraag *aan bod komt* (*mentioned*), of dat *expliciet of impliciet* gebeurt, met welk *sentiment* en als afsluiting een kort antwoord van twee tot vier zinnen. Het volgende werkwijzedeel uit de prompt laat zien hoe het model zijn redenering heeft opgebouwd.

“Per call krijg je ÉÉN onderzoeksvraag en ÉÉN verhaal. Je beantwoordt de vraag uitsluitend op basis van wat in het verhaal staat – geen verzinsels, geen externe kennis, geen aannames.”

Werkwijze (twee-turn chain-of-thought in één gestructureerd antwoord)

1. Lees de onderzoeksvraag en de verwachte signalen aandachtig.
2. Lees het verhaal grondig.
3. Schrijf eerst kort je REDENERING (max 60 woorden, Nederlands) waarin je identificeert welke fragmenten van het verhaal relevant zijn voor de vraag.
4. Citeer letterlijk: geef ALLE relevante passages (TOT 5), elk 1-3 zinnen, met per quote het exacte source_field. Quotes mogen elkaar niet overlappen.
5. Bepaal mentioned, type, sentiment op basis van de quotes.
6. Beoordeel prominence en latent - verplicht wanneer mentioned=true.
7. Schrijf een ANSWER (2-4 zinnen Nederlands): hoe adresseert dit verhaal deze onderzoeksvraag? Vat samen wat de vrouw hierover vertelt, gegrond In de geciteerde passages.

2. Inductief coderen

Naast de deductieve score haalde het model in een tweede aanroep nul tot acht *codes* uit hetzelfde verhaal die buiten de registry vallen. Deze codes relateerden aan de overkoepelende onderzoeksvragen van dit project.

Zo werden geen patronen gemist die vooraf niet bedacht waren. Het model schreef eerst een korte samenvatting van het verhaal en formuleerde daarna codes met een Nederlandse label, één tot vijf letterlijke passages, en een inschatting van *prominence* en latent. Het kader uit de prompt:

Werkwijze

1. Lees het verhaal grondig.
2. Schrijf een korte samenvatting (≤80 woorden Nederlands).
3. Identificeer 3-8 codes. Elke code:
   - heeft een korte Nederlandse label (2-5 woorden)
   - heeft 1 TOT 5 letterlijke passages uit het verhaal (1-3 zinnen) met source_field - pak alle passages die deze code ondersteunen
   - is NIET te generiek ("vermoeidheid", "frustratie") en NIET te idiosyncratisch ("siliconen implantaten van een specifiek merk")
   - is relevant voor de werkagenda
   - krijgt prominence + latent toegekend (verplicht).

Codes kwamen niet uit een vooraf opgestelde lijst - ze emergeerden uit het verhaal. Citeer letterlijk; hallucineer niet.

3. Consolideren (scores naar thema's)

Eén corpus-brede LLM-aanroep voegde de circa 6.900 ruwe codes samen tot 36 thema's: 20 inductief en 16 deductief (één thema per registry-vraag). Per thema produceerde het model een korte definitie, een overkoepelend concept, voorbeeldcitaten en het aantal verhalen waarin het thema voorkwam.

4. Rescoren

De eerste consolidatie leverde thema's op basis van de oorspronkelijke per-verhaal-codes. Om frequenties zuiver te maken, doorliep het model elk verhaal opnieuw en bepaalde het per thema of dat thema in dat verhaal aanwezig was. Hiermee werd ervoor gecorrigeerd dat de telling voor verhalen die een thema raakten maar in het inductieve coderen niet als losse code werd verkregen.

5. Groeperen tot overkoepelende thema's

De 20 inductieve thema's werden via embedding-clustering en een afsluitende LLM-naam-aanroep gegroepeerd tot 16 *overkoepelende thema's*. Deze grovere laag bood een hoofdstructuur voor het rapport zonder de specificiteit van de onderliggende thema's te verliezen.

6. Synthese per onderzoeksvraag

Tot slot schreef het model per onderzoeksvraag een lopende samenvatting, opgebouwd uit de geciteerde passages en de gestructureerde scores. Sinds deze versie produceerde de pijplijn óók een gestratificeerde variant per leeftijdsgroep, zodat leeftijdsverschillen direct zichtbaar werden.

3.2.3 Vocabulaire en literatuur-anker

Het vocabulaire volgde de hiërarchie uit de reflexieve thematische analyse (Braun & Clarke, 2022): een *code* is één observatie in één verhaal, een *thema* een patroon over het hele corpus, en een *overkoepelend thema* een cluster van samenhangende thema's. Per thema formuleerde het model een overkoepelend concept (in de Engelstalige literatuur *central organising concept*; Braun & Clarke, 2022), te lezen als de kerngedachte die alle codes binnen dat thema samenbindt.

De epistemologie sloot aan bij de meer gestructureerde varianten van thematische analyse: framework analyse (Gale et al, 2013) en thematische inhoudsanalyse (Boyatzis, 1998), omdat er frequenties werden geteld en er met een vooraf opgesteld codeboek werd gewerkt. De vraagregistry was theoretisch verankerd via herkomstlabels die verwijzen naar constructen uit de literatuur: epistemische onrechtvaardigheid (Kidd & Carel, 2017), diagnostische gender-bias en psychologisering (Harrilal-Maharaj, 2025), en de taxonomie van *unheard mechanisms* uit het NIHR-rapport *Feeling unheard* (National Institute for Health and Care Research, 2024).

3.2.4 Mens-in-de-lus voor validatie

De analytische pijplijn is bewust niet volautomatisch; mens-in-de-lus-validatie scherpt het codeboek aan en toetst de modelscores. Twee onderzoekers codeerden onafhankelijk dezelfde steekproef uit de kalibratierun. Merlijn van Breugel codeerde 25 verhalen op alle 16 registryvragen (400 verhaal-vraagparen); Hester Verhoeven, basisarts, codeerde 10 overlappende verhalen (159 gedeelde cellen). Op themapresentie (komt de vraag aan bod) was de interbeoordelaarsbetrouwbaarheid substantieel: Cohen's κ = 0,76 bij 88% ruwe overeenstemming (Landis & Koch: 0,61–0,80 geldt als substantieel). De LLM kwam met elke onderzoeker ongeveer even goed overeen als de onderzoekers onderling (κ = 0,67 met Merlijn, 0,74 met Hester), wat aangeeft dat het model binnen de normale marge van menselijke codeervariatie opereert en de afgesproken drempel van κ ≥ 0,65 haalt. Voor het secundaire sentimentoordeel ligt de ruwe overeenstemming hoog (85-96%), maar valt de κ lager uit (0,35-0,65); dat is een kappa-paradox-artefact, doordat het sentiment in dit corpus overwegend negatief is en er weinig variatie is waarop κ kan belonen. Op het inductieve spoor bevestigde een aanvullende beoordeling van 101 verhaal-themaparen over 7 verhalen de themapresentie in 100% van de gevallen. Deze validatie is uitgevoerd op de kalibratierun; de hieruit volgende aanscherpingen (scherpere scope-notes en scoringsrubriek) zijn in de definitieve run verwerkt (zie ook Bijlage C voor een methodologische reflectie).

4 Resultaten

4.1 Resultaten van de kwantitatieve analyse

De gemiddelde leeftijd van de 1637 respondenten was 50,5 jaar (range 15-86); de gemiddelde leeftijd waarop hun gezondheidsklachten begonnen was 33,5 jaar (range 0-76 jaar).

4.1.1 Ervaren klachten

Tabel 3 *Aantal gerapporteerde lichamelijke klachten uit beide enquêtes (n=1637)*

Welke (onverklaarde) lichamelijke klachten ervaar(de) je?	***Enquête-2 + Enquête-3*** **n=**	**Percentage %**
Allergieën	427	26
Buikpijn	511	31
Droge huid	475	29
Duizeligheid	636	39
Gebitsproblemen	243	15
Geheugenverlies	598	37
Gewichtstoename	674	41
Haaruitval	465	28
Hoofdpijn	669	41
Hyperventilatie	220	13
Incontinentie	208	13
Infecties	271	17
Klachten spijsvertering	**721**	**44**
Menstruatieproblemen	567	35
Misselijkheid	379	23
Nachtelijk zweten	565	35
Onregelmatige hartslag	568	35
Onregelmatige menstruatie	347	21
Ontstekingen	443	27
Onverklaarbare pijn	**697**	**43**
Oorsuizen	443	27
Opgeblazen gevoel	672	41
Opvliegers	474	29
Pijnlijke gewrichten	**940**	**57**
Spierpijn	**704**	**43**
Spierspanning	668	41
Verminderd libido	624	38
Vermoeidheid	**1259**	**77**

Vetgedrukt de top 5 van meest gerapporteerde lichamelijke gezondheidsklachten: vermoeidheid (ervaren door 77% van de vrouwen die een van de enquêtes heeft ingevuld); pijnlijke gewrichten (57%); spijsverteringsklachten (44%); onverklaarde pijn (43%); en spierpijn (43%). Het minst frequent werden gerapporteerd: gebitsproblemen, hyperventilatie en incontinentie.

Tabel 4 *Aantal gerapporteerde psychische klachten uit beide enquêtes (n=1637)*

Welke (onverklaarde) psychische klachten ervaar(de) je?	*Enquête-2 + Enquête-3* n	Percentage %
Angsten	687	42
Apathie	183	11
Concentratieverlies	**856**	**52**
Depressief	615	38
Eenzaamheid	508	31
Eetproblemen	302	18
Gedachten aan zelfdoding	409	25
Gevoel van falen	707	43
Hoogsensitiviteit	652	40
Nervositeit	603	37
Onzekerheid	673	41
Paniekaanvallen	528	32
Piekeren	722	44
Prestatiedruk	426	26
Prikkelbaarheid	691	42
Rouw	385	24
Slaapproblemen	**953**	**58**
Somberheid	**794**	**49**
Stemmingswisselingen	**833**	**51**
Stress	**805**	**49**
Verdriet	549	34
Verwardheid	281	17

De top 5 van meest gerapporteerde psychische gezondheidsklachten: slaapproblemen (58%); concentratieverlies (52%); stemmingswisselingen (51%); somberheid (49%); en, stress (49%). Het minst vaak gerapporteerd werden: eetproblemen (18%), verwardheid (17%) en apathie (11%).

4.1.2 Gestelde diagnosen

De vraag naar welke diagnose er gesteld was naar aanleiding van de (onverklaarde) gezondheidsklachten werd als volgt beantwoord:

Tabel 5 *Gestelde diagnosen*

Diagnose	n	%
Auto-immuunziekte	151	12
Angststoornis / hyperventilatie	196	15
Burn-out / stress	**434**	**34**
Depressie	**303**	**24**
Endometriose	100	8
Fibromyalgie	185	15
Hart- en vaatziekten	79	6
Menstruatiestoornis	113	9
Migraine	159	12
Overgang	**316**	**25**
PCOS	45	4
Schildklieraandoening	129	10
Geen diagnose gesteld	300	24

Burn-out/stress, de overgang en depressie waren de meest gestelde diagnosen bij respectievelijk 34%, 25% en 24% van de vrouwen die de enquête hadden ingevuld.

Ook is de vraag gesteld wat volgens hen uiteindelijk de juiste diagnose bleek te zijn (tabel 6). Deze en de volgende vragen zijn maar door een deel van de respondenten ingevuld.

Tabel 6 *Uiteindelijk juiste diagnose*

Diagnose	**n**	**%**
Auto-immuunziekte	**57**	**4**
Angststoornis / hyperventilatie	8	1
Burn-out / stress	15	1
Depressie	16	1
Endometriose	**43**	**3**
Fibromyalgie	22	2
Hart- en vaatziekten	16	1
Menstruatiestoornis	19	1
Migraine	21	2
Overgang	**140**	**11**
PCOS	11	1
Schildklieraandoening	40	3
Weet ik niet	135	11

Als antwoord op de vraag naar welke diagnose uiteindelijk is gesteld werd het meest vaak de overgang en een auto-immuunziekte genoemd.

In de enquête is ook de vraag voorgelegd of men dacht dat de klachten verband hielden met menstruatie, zwangerschap of de overgang (tabel 7).

Tabel 7 *Relatie met menstruatie, zwangerschap of de overgang*

Ervaren verband	**n**	**%**
Verband met menstruatie	498	30
Verband met zwangerschap	247	15
Verband klacht met de overgang	766	47
Geen verband	305	19
Weet ik niet	298	18

Ongeveer de helft van de respondenten dacht zelf dat hun klachten verband hielden met de overgang, ook de menstruatie werd als belangrijke oorzaak genoemd.

Tabel 8 *Verwijzingen naar specialist*

Specialist	n	%
Anesthesist	85	5
Cardioloog	358	22
Dermatoloog	210	13
Gynaecoloog	693	42
Internist	469	29
Neuroloog	419	26
Oncoloog	27	2
Orthopeed	202	12
Psycholoog	631	39
Reumatoloog	359	22

De meeste respondenten die verwezen waren werden verwezen naar de gynaecoloog (42%), psycholoog (39%) of internist (29%).

4.1.3 Klachten en diagnosen naar leeftijd

Tabel 9 *Top 5 klachten en diagnosen voor de twee grootste leeftijdscategorieën*

Lichamelijke klachten		Diagnose	
25-49 jaar	*50-65 jaar*	*25-49 jaar*	*50-65 jaar*
Vermoeidheid	Vermoeidheid	Burn-out/stress	Overgang
Pijnlijke gewrichten	Pijnlijke gewrichten	Geen diagnose	Burn-out/stress
Hoofdpijn	Spierpijn	Depressie	Depressie
Opgeblazen gevoel	Gewichtstoename	Angststoornis	Schildklieraandoening
Spijsverteringsklachten	Verminderd libido	Overgang	Fibromyalgie

Psychische klachten	
25-49 jaar	*50-65 jaar*
Slaapproblemen	Slaapproblemen
Stemmingswisselingen	Concentratieverlies
Stress	Stemmingswisselingen
Somberheid	Stress
Concentratieverlies	Somberheid

In de top 5 van lichamelijke klachten kwamen vermoeidheid en pijnlijke gewrichten het meest frequent voor in de twee grootste leeftijdsgroepen; hoofdpijn blijkt vaker genoemd bij de jongere leeftijdsgroep en verminder libido bij de oudere leeftijdsgroep. De diagnose angststoornis werd vaker genoemd door jongere vrouwen en de diagnosen schildklieraandoening en fibromyalgie juist door oudere vrouwen. Bij de psychische klachten werden inhoudelijk geen verschillen gevonden tussen de leeftijdsgroepen.

4.1.4 Klachten en diagnosen naar klachtenduur

Tabel 10 *Top 5 klachten en diagnosen naar duur klachten*

Lichamelijke klachten		**Diagnose**	
0-9 jaar	*10 jaar of langer*	*0-9 jaar*	*10 jaar of langer*
Vermoeidheid	Vermoeidheid	Geen diagnose	Burn-out/stress
Pijnlijke gewrichten	Pijnlijke gewrichten	Burn-out/stress	Depressie
Geheugenverlies	Spijsverteringsklachten	Overgang	Overgang
Spierspanning	Onverklaarde pijn	Depressie	Angststoornis
Gewichtstoename	Spierpijn	Angststoornis	Migraine

Psychische klachten	
0-9 jaar	*10 jaar of langer*
Concentratieverlies	Slaapproblemen
Stemmingswisselingen	Concentratieverlies
Slaapproblemen	Somberheid
Stress	Stemmingswisselingen
Somberheid	Stress

De klachtenduur maakte weinig uit voor het type klachten of de gestelde diagnose.

4.1.5 Gehoord voelen

Tabel 11 *Aantal vrouwen dat zich gehoord voel(d)en bij de arts*

Gehoord voelen	**n**	**%**
Bij specialist	337	24
Bij huisarts	356	22

Uit de enquêtes komt naar voren dat - over het geheel genomen - ongeveer 1 op de 4 respondenten zich gehoord voelen bij de huisarts en specialist. Deze aantallen verschilden nauwelijks voor de twee grootste leeftijdsgroepen en de groepen die korter of langer met hun klachten rondliepen.

4.1.6 Impact van de klachten op levensdomeinen

Tabel 12 *Aantal vrouwen dat impact op persoonlijk leven ervaarden*

Impact	**n**	**%**
Gezin/relatie	1489	92
Sociale contacten	1461	90
Financieel	1012	63

Bijna alle vrouwen gaven aan dat hun (onverklaarde) gezondheidsklachten invloed hadden op hun persoonlijke relaties en veel van hen rapporteerden ook financiële consequenties.

4.1.7 Validatie resultaten enquêtes

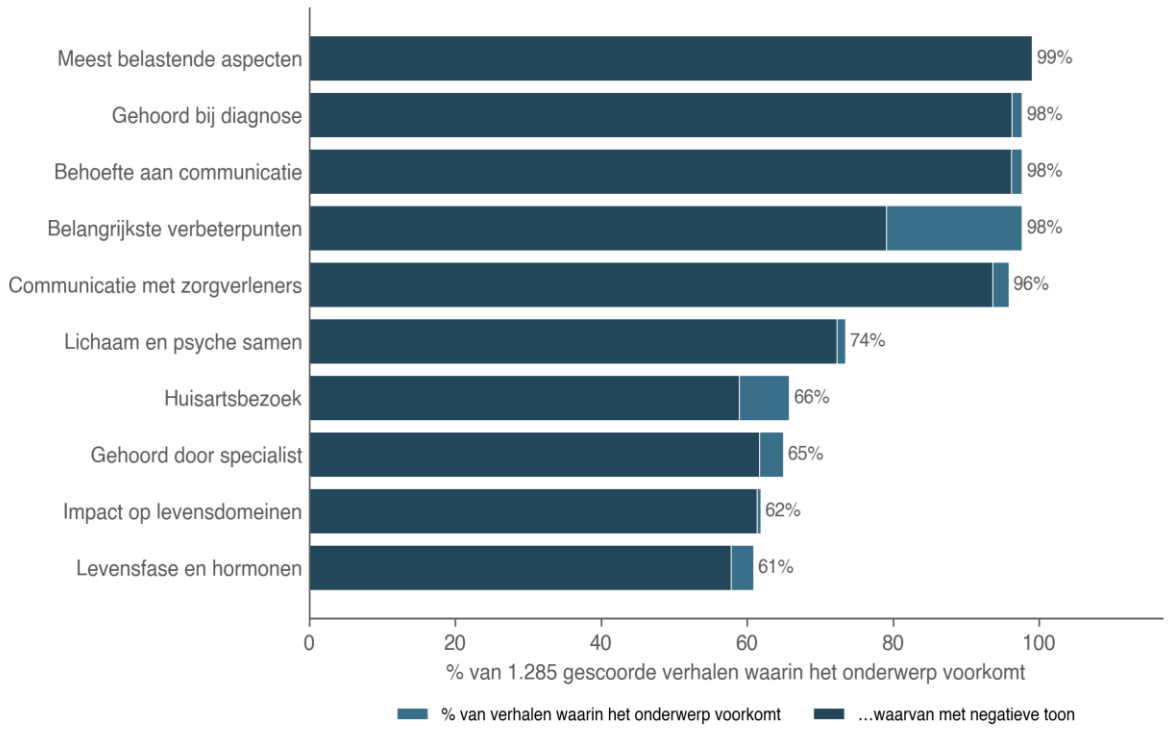
De resultaten uit *Enquête-1*, ingevuld door 2268 vrouwen, zijn waar mogelijk vergeleken met die uit de meer recent ingevulde enquêtes. Daaruit bleek dat meer vrouwen die *Enquête-1* hadden ingevuld tot de leeftijdscategorie 25-49 jaar behoorden (53% tegenover 38%) en minder tot de groep 50-65 jaar (37% tegenover 50%). Ook bleek een klachtenduur van 1-5 jaar voor meer vrouwen uit *Enquête-1* (33%) te gelden dan voor vrouwen uit de meer recente enquête (21%). In beide enquêtes ervaarden de meeste respondenten 5 jaar of langer klachten (*Enquête-1* 74%; *Enquête-3* 62%). In *Enquête-1* rapporteerden meer vrouwen dat ze geen diagnose hadden gekregen (39%) dan in de *Enquête-3* (18%). Tenslotte is gekeken naar de antwoorden op de vragen of men zich gehoord voelde bij de arts. In *Enquête-1* reageerden men positiever op de vraag of men zich gehoord voelde bij de huisarts (33% voelde zich gehoord) dan in *Enquête-3* (22%). Hetzelfde beeld kwam naar voren in de antwoorden op de vraag betreffende je gehoord voelen bij de specialist (34% versus 24%). Ondanks deze verschillen lijkt het beeld overall vergelijkbaar tussen de verschillende metingen: de meeste respondenten hebben al jarenlang gezondheidsklachten en voelen zich niet gehoord door huisarts of specialist.

4.2 Resultaten van de kwalitatieve analyse

4.2.1 Deductieve spoor

Het deductieve spoor heeft alle vragen op alle 1.285 verhalen gescoord (20.560 verhaal-vraag-paren, 0 fouten). Figuur 1 toont de tien meest voorkomende onderzoeksvragen, met daarbinnen het aandeel verhalen met een negatieve toon. De vijf grootste onderwerpen bleken vrijwel universeel: zij kwamen in 96% tot 99% van de verhalen voor. Dat was op zichzelf een bevinding. Het meldpunt trekt vrouwen aan met een gedeelde kern van ervaringen, en die kern draait niet om een specifieke aandoening maar om erkenning, communicatie en de weg naar een diagnose.

Figuur 1 *Top-tien onderzoeksvragen uit de registry, geordend naar prevalentie. De donkere balk binnen elke balk geeft het aandeel met negatieve toon weer*



Hieronder volgt per onderzoeksvraag een gesynthetiseerd antwoord van enkele zinnen, met de meest voorkomende deelclusters (de “sub-thema’s” die de pipeline binnen die vraag identificeerde). De vragen zijn geordend volgens de zes categorieën uit het Plan van Aanpak.

Categorie 1. Aard en duur van klachten

1. Hoe wordt de samenhang tussen lichamelijke en psychische klachten ervaren én geïnterpreteerd door behandelaars? (53,5%, n=687) Vrouwen beschrijven een onlosmakelijke wisselwerking tussen lichaam en geest: hormonale schommelingen, chronische pijn en uitputting raken direct het mentale welzijn, en omgekeerd versterken stress en trauma de fysieke kwetsbaarheid. Behandelaars interpreteren die samenhang doorgaans eenzijdig en grijpen, zodra een directe medische verklaring ontbreekt, naar een psychische duiding. In ruim de helft van de verhalen (53,5%) krijgt die psychische duiding een concrete vorm: vrouwen worden naar huis gestuurd met het advies rust te nemen, te gaan sporten of een psycholoog te bezoeken. De clusters: *Stress en burn-out als standaarddiagnose* (n=140), *Het "tussen de oren" stempel* (n=128), *Onterechte psychiatrische diagnoses en medicatie* (n=141), *Dossier-bias en eerdere geschiedenis* (n=26) en *Reduceren tot levenssituatie of moederschap* (n=31).

2. Welke patronen in klachten komen naar voren in relatie tot vrouwelijke levensfasen en hormonale veranderingen? (60,9%, n=783) Hormonaal-bepaalde levensfasen vormen een rode draad: puberteit, menstruatie, zwangerschap en vooral de (vroege) overgang zetten klachten in gang of versterken bestaande klachten. Vrouwen ervaren dat deze hormonale samenhang structureel wordt onderschat. De clusters: *impact van de overgang* (n=15), *cyclus- en menstruatie-gerelateerde klachten* (n=12), *zwangerschap en postpartum-ontregeling* (n=10), *hormonale misdiagnoses en bagatellisering* (n=14) en een expliciete *behoefte aan onderzoek en erkenning* (n=8).

Categorie 2. Diagnoseroute en erkenning

3. Hoe ervaren vrouwen het proces van diagnostiek en de rol van zorgverleners? Voelen zij zich gehoord en serieus genomen? (97,7%, n=1.256) Het diagnostisch proces wordt overwegend ervaren als een moeizame, langdurige en eenzame zoektocht. Klachten worden gebagatelliseerd, afgedaan als “normale vrouwenkwaaltjes” of direct toegeschreven aan psychische oorzaken; vertragingen met soms ernstige gevolgen zijn het resultaat. Het gebrek aan gehoor leidt ertoe dat vrouwen aan hun eigen waarneming gaan twijfelen. De clusters: *psychologisering van fysieke klachten* (n=22), *bagatelliseren en “het hoort erbij”* (n=15), *strijden voor een luisterend oor* (n=14), *zelf twijfelen en zorgmijding* (n=12) en een *behoefte aan een holistische blik* (n=10).

4. Welke rol spelen psychische labels (burn-out, depressie, angststoornis) in het traject naar een lichamelijke diagnose? (57,4%, n=738) Psychische labels werken als een filter dat verder somatisch onderzoek belemmert. Eenmaal in het dossier opgenomen, kleuren zij elk vervolgcontact: nieuwe fysieke klachten worden door diezelfde bril bekeken en niet of vertraagd onderzocht. De clusters: *fysieke aandoeningen afgedaan als stress* (n=26), *het ‘tussen de oren’-stigma* (n=25), *zelftwijfel en zorgmijding door psychologisering* (n=9) en *diagnostische overschaduwing door bestaande labels* (n=7).

Categorie 3. Zorggebruik en zorgprofessionals

5. Hoe beschrijven vrouwen hun contactmomenten met de huisarts en de ervaren kwaliteit daarvan? (48,2%, n=619) De huisarts wordt vaak als knelpunt ervaren. Chronische pijn, vermoeidheid en hormonale disbalans worden herhaaldelijk afgedaan als stress, hyperventilatie of burn-out; doorverwijzing moet bevochten worden. Tegelijk benoemen vrouwen ook ervaringen met huisartsen die wel doortastend handelen. Doordat lichamelijk of bloedonderzoek daardoor achterwege blijft,

lopen diagnoses in veel gevallen jarenlange vertraging op. De clusters: *Psychologisering van fysieke klachten* (n=130), *Bagatelliseren van gendersensitieve klachten* (n=65), *Barrières bij doorverwijzingen* (n=101), *Gewichtsstigma en vooroordelen* (n=9) en *Erkenning en daadkracht als uitzondering* (n=57).

6. In hoeverre voelen vrouwen zich gehoord door specialisten, en verschilt dit van de huisartsen-ervaring? (77,7%, n=999) Bij specialisten lopen vrouwen tegen een vergelijkbare barrière aan als bij de huisarts, met meer variatie. Soms is juist de specialist degene die de klacht alsnog serieus neemt nadat de huisarts deze afdeed, soms is het omgekeerd; de rode draad is de hokjesgeest van het specialisme. Zodra standaardonderzoek niets aantoont, voelen vrouwen zich ook bij de specialist al snel niet geloofd, wat gevoelens van machteloosheid versterkt. De clusters: *Psychologisering van fysieke klachten* (n=294), *Gebrek aan een holistische blik* (n=109), *Van het kastje naar de muur* (n=49), *Contrast tussen huisarts en specialist* (n=83) en *Geringe empathie en bejegening* (n=288).

7. Hoe beschrijven vrouwen de communicatie met zorgverleners, en welke elementen dragen bij aan positieve of negatieve ervaringen? (95,9%, n=1.232) De communicatie wordt overwegend als moeizaam en kleinerend beschreven. Verschillende vrouwen benoemen het fenomeen expliciet als *medical gaslighting*, waarmee een zorgverlener (onbedoeld) een patiënt het gevoel geeft dat de klachten zijn ingebeeld. Het stelselmatig weggeredeneerd worden ondermijnt vertrouwen en kan leiden tot zorgmijding. Positieve ervaringen zijn zeldzaam en hangen vrijwel altijd samen met een arts die de tijd neemt en doorvraagt. De clusters: *luisteren als sleutel tot herstel* (n=25), *psychologisering van fysieke klachten* (n=22), *bagatelliseren en kleineren* (n=18), *strijden voor gehoor en doorverwijzing* (n=15) en *behoefte aan een holistische benadering* (n=12).

8. Hoe ervaren vrouwen van kleur hun interacties met zorgverleners, en welke verschillen zien zij ten opzichte van andere vrouwen? (1,2%, n=15) Het zeer lage percentage weerspiegelt de samenstelling van het corpus, niet de afwezigheid van het probleem. De weinige verhalen die dit thema raken beschrijven specifieke barrières: somatische tekens worden niet herkend op een donkere huid, etnische stereotypen leiden tot bagatellisering, en relevante genetische risico's worden niet onderkend. De clusters: *gebrek aan representatieve medische kennis* (n=6), *etnische stereotypen en vooroordelen* (n=4), *onherkenbaarheid van fysieke symptomen* (n=3) en *communicatiebarrières en uitsluiting* (n=2). Gerichte aanvullende dataverzameling is noodzakelijk om hier methodologisch onderbouwde uitspraken over te doen.

Categorie 4. Behandeling en medicatie

9. In welke mate wordt rekening gehouden met de hormoonhuishouding bij het voorschrijven van medicatie? (24,0%, n=308) Hormonale schommelingen en cyclusafhankelijke effecten worden in de praktijk vaak genegeerd. In plaats van gericht hormonaal onderzoek volgt eenzijdige symptoombestrijding, met de anticonceptiepil als veelvoorkomend standaardantwoord. De clusters: *gebrek aan hormoon-bewustzijn* (n=22), *de pil als standaardoplossing* (n=18), *psychologisering en symptoombestrijding* (n=15), *negeren van patiëntintuïtie* (n=12) en een *roep om gendersensitieve medicatie* (n=10).

10. Hoe wordt omgegaan met meldingen van bijwerkingen, en hoe komen vrouwen tot het besluit om met medicatie te stoppen? (19,0%, n=244) Bijwerkingen worden vaak gebagatelliseerd of als "hinderlijk" weggezet; in plaats van dosisaanpassing volgt soms een extra medicament tegen de bijwerking. Een aanzienlijk deel van de vrouwen besluit daarom autonoom te stoppen. De clusters: *autonoom stoppen wegens ernstige klachten* (n=15), *wegwimpelen en bagatelliseren van bijwerkingen* (n=12), *dwang en onbegrip bij weigering* (n=8), *symptoombestrijding met extra medicatie* (n=4) en *medicatie afgestemd op de man* (n=4).

Categorie 5. Maatschappelijke en persoonlijke impact

11. Op welke manieren beïnvloeden klachten verschillende levensdomeinen (gezin, sociale contacten, financiën)? (77,4%, n=995) De impact is diepgaand en raakt vrijwel elk levensdomein. Vrouwen verliezen sociale contacten, kunnen hun ouderrol of werk niet meer waarmaken en belanden in isolement; dit isolement wordt versterkt door breed gedeeld onbegrip. Op het werk leidt dit vaak tot gedwongen parttime werken of arbeidsongeschiktheid, waarbij sommige vrouwen hun ervaringen met betrokken instanties als traumatiserend beschrijven. De clusters: *Verlies van werk en carrière* (n=240), *Beperkingen in het dagelijks functioneren* (n=636), *Sociaal isolement en onbegrip* (n=192), *Druk op relaties en gezinsleven* (n=219) en *Mentale uitputting en angst* (n=214).

12. Wat beschrijven vrouwen als de meest belastende aspecten van hun zorg- en ziekteproces? (99,1%, n=1.274) Het meest belastend is niet de ziekte zelf maar het structurele gebrek aan erkenning. Vrouwen beschrijven dat hun klachten worden gebagatelliseerd of gepsychologiseerd, wat leidt tot zelftwijfel en het gevoel er alleen voor te staan; daar bovenop komt de fysieke en mentale uitputting van een jarenlange diagnostische zoektocht. De clusters: *niet serieus genomen worden* (n=32), *langdurige onzekerheid en diagnosetraject* (n=25), *verlies van functioneren en isolement* (n=22), *onterechte psychologisering* (n=18) en *zelftwijfel en angst* (n=15).

Categorie 6. Behoeften en verbeterpunten

13. Wat zijn de meest genoemde verbeterpunten in de zorg voor vrouwen met onverklaarde klachten? (97,7%, n=1.255) Het meest urgente verbeterpunt is dat artsen écht luisteren en eerst lichamelijke oorzaken grondig uitsluiten voordat een psychische verklaring in beeld komt. Daarnaast pleiten vrouwen voor meer onderzoek naar het vrouwenlichaam, integrale zorg en snellere diagnostiek. De clusters: *echt luisteren en serieus nemen* (n=35), *gendersensitieve kennis en onderzoek* (n=25), *snellere diagnostiek en betere organisatie* (n=20), *stoppen met onterecht psychologiseren* (n=18) en *holistische en integrale benadering* (n=12).

14. Welke behoeften noemen vrouwen als het gaat om communicatie, erkenning en begeleiding? (97,7%, n=1.255) Vrouwen vragen om zorgverleners die echt luisteren, doorvragen en de expertise van de patiënt over haar eigen lichaam respecteren. Daarnaast is er een sterke roep om eerst lichamelijke oorzaken uit te sluiten en om brede, integrale aandacht. De clusters: *echt luisteren en serieus nemen* (n=42), *fysieke oorzaken uitsluiten* (n=24), *gendersensitieve kennis en scholing* (n=16), *holistische en integrale blik* (n=15) en *snellere en betere doorverwijzing* (n=12).

15. Welke informatiebehoeften hebben vrouwen (diagnoses, hormoonwerking, behandelopties)? (47,8%, n=614) Vrouwen ervaren een structureel tekort aan betrouwbare, gendersensitieve informatie over hun eigen lichaam. Dit dwingt hen tot eigen onderzoek via internet, fora en alternatieve kanalen, met wisselende kwaliteit. De grootste behoefte betreft inzicht in de onderliggende oorzaak en samenhang van klachten, de werking van hormonen en de beschikbare behandelopties met hun bijwerkingen. De clusters: *Oorzaken en samenhang van klachten* (n=243), *Werking en impact van hormonen* (n=186), *Behandelopties en bijwerkingen* (n=158), *Gendersensitieve gezondheidsinformatie* (n=92) en *Navigatie en geschikte zorgpaden* (n=26).

16. Welke rol spelen lotgenotencontact en zelfhulp in het omgaan met onverklaarde klachten? (25,7%, n=331) Lotgenotencontact en zelfhulp vullen actief de leemtes die de reguliere zorg achterlaat. Vrouwen “dokteren” massaal zelf, lezen medische literatuur, vragen testen aan en wisselen kennis uit; dit wordt zwaar gevonden maar ook validerend. Lotgenotencontact is een van de weinige zorgthema’s waar vrouwen óók met tevredenheid over spreken. De clusters: *gedwongen zelf*

dokteren (n=52), *validatie en herkenning door lotgenoten* (n=22), *kennisdeling via informele netwerken* (n=15) en, als kritische tegenstem, *wantrouwen jegens zelfonderzoek* (n=5).

4.2.2 Profiel per domein en leeftijdsgroep

Domein: medisch, emotioneel of sociaal. Per geconstateerde aanwezigheid van een thema kent de pijplijn drie aanvullende labels toe die de aard van de last beschrijven: of die *medisch*, *emotioneel* of *sociaal* van karakter is, en of dat domein in dat verhaal *centraal*, *bijkomend* of *afwezig* is. Over het hele corpus is het medische domein veruit het vaakst centraal aanwezig (70,9% van de waarnemingen), gevolgd door het emotionele (18,3% centraal, 45,7% bijkomend) en het sociale (10,4% centraal, 78,1% afwezig). De medische lens domineert dus de manier waarop vrouwen hun ervaring vertellen, maar emotionele last loopt vrijwel overal mee als secundaire dimensie.

Per thema viel het accent duidelijk anders. De grote inductieve thema's *decennialange diagnostische vertraging*, *symptoombestrijding zonder bronaanpak*, *huisarts als poortwachtersbarrière* en *hormonale triggers en overgangsklachten als blinde vlek* waren vrijwel zuiver medisch (98–100% centraal medisch). Daarentegen was *zelftwijfel en verlies van lichaamsvertrouwen* sterk emotioneel (100%), en *verlies van identiteit, levenskwaliteit en werk* tegelijk sterk emotioneel én sociaal (67% emotioneel, 64% sociaal centraal). *Sociaal isolement en maatschappelijk onbegrip* was het meest uitgesproken sociale thema (92% centraal sociaal). Het mechanisme *niet serieus genomen worden* was medisch én emotioneel zwaar (89% en 74% centraal), wat aansluit op de eerder beschreven *medical gaslighting*-dynamiek.

Leeftijd. Leeftijd was bekend voor 578 van de 1.285 verhalen; de jongste groep (10–25 jaar) telde slechts 21 verhalen en moet daarom voorzichtig gelezen worden. De gestratificeerde synthese liet zien dat de top-vijf onderzoeksvragen leeftijdsonafhankelijk universeel zijn (95% tot 100% in alle groepen), terwijl de inhoud verschuift met de levensfase. De samenhang tussen lichamelijke en psychische klachten woog sterker bij jongere vrouwen (81% bij 10–25, 70% bij 55+). De huisarts als knelpunt was opvallend prominent in de jongste groep (81% versus 60% bij 55+), wat past bij de drempel die jonge vrouwen ervaren om door deze poortwachter heen te komen. Onderwerpen rond hormonen en levensfasen liepen logischerwijs hoger op in de middengroep en bij ouderen, met de overgang als markant snijvlak. De domeinverhouding (medisch, emotioneel, sociaal) was over de leeftijdsgroepen opvallend stabiel: wélke thema's spelen verschilde, maar de samenstelling van de last verschoof niet wezenlijk.

4.2.3 Inductieve bevindingen

Naast het deductieve scoren produceert het inductieve spoor open codes die buiten de registry (de vooropgestelde deelvragen c.q. thema's) vallen. Consolidatie reduceerde ongeveer 6.900 ruwe codes tot twintig inductieve thema's; een rescoring-pas berekende vervolgens voor elk verhaal opnieuw of het thema aanwezig is. Figuur 2 toont de twaalf meest voorkomende thema's na rescoring. De top vier komt voor in meer dan 60% van de verhalen en bevestigt langs een onafhankelijke route de centrale ervaringspatronen uit het deductieve spoor: niet serieus genomen worden (n=1.115), decennialange diagnostische vertraging (n=936), verlies van identiteit, levenskwaliteit en werk (n=817) en de overweldigende last van onverklaarde, multi-systemische klachten (n=807). De rescoring slaagde voor 1.265 van de 1.285 verhalen; twintig verhalen vielen uit door tijdelijke API-time-outs en blijven in de themafrequenties buiten beschouwing.

Figuur 2 *De twaalf meest aanwezige inductieve thema's na rescoring. De top vier komt voor in meer dan 60% van de gescoorde verhalen*



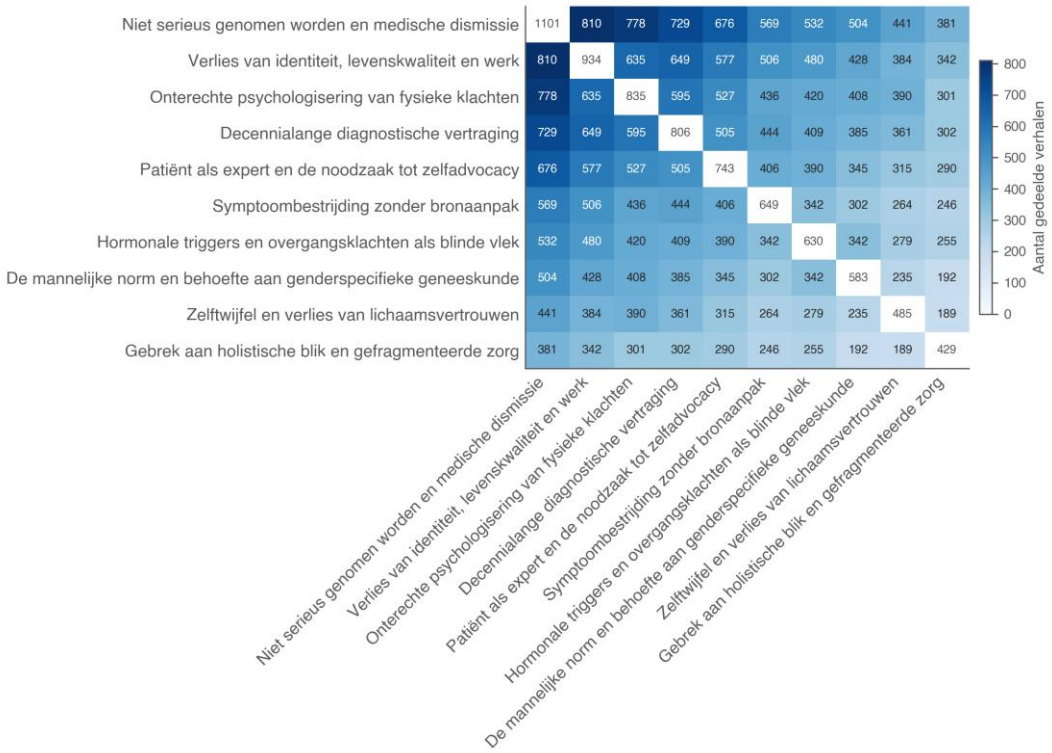
Tabel 13 *Top-tien inductieve thema's na rescoring (van 1.285 verhalen), met overkoepelend concept*

Thema	Overkoepelend concept (samenvatting)	n
Niet serieus genomen worden en medische dismissie	Het stelselmatig wegredeneren of bagatelliseren van klachten ondermijnt erkenning en vertrouwen.	1.101
Verlies van identiteit, levenskwaliteit en werk	Langdurige klachten leiden tot baanverlies, sociaal isolement en het wegvallen van toekomstperspectief.	934
Onterechte psychologisering van fysieke klachten	Lichamelijke symptomen worden toegeschreven aan stress, depressie of trauma zonder differentiële diagnostiek.	835
Decennialange diagnostische vertraging	Een jarenlang traject van verkeerde of uitblijvende diagnoses voorafgaand aan de juiste herkenning.	806
Patiënt als expert en de noodzaak tot zelfadvocacy	Vrouwen worden gedwongen zelf de regie, diagnostiek en doorverwijzing te organiseren.	743
Symptoombestrijding zonder bronaanpak	Een behandelroute beperkt tot pijnstilling of hormoonpreparaten, zonder onderliggend onderzoek naar oorzaak.	649
Hormonale triggers en overgangsklachten	De hormonale dimensie wordt structureel onderschat in zowel diagnose als behandeling.	630
De mannelijke norm en gendersensitieve geneeskunde	Protocollen en referentiewaarden zijn op mannen geijkt; gendersensitieve patronen worden gemist.	583
Zelftwijfel en verlies van lichaamsvertrouwen	Het verinnerlijken van afwijzing leidt tot wantrouwen in het eigen lichaam en uitstel van nieuwe hulpvragen.	485
Gebrek aan holistische blik en gefragmenteerde zorg	Symptomen worden los van elkaar behandeld; samenhang en context blijven onbenut.	429

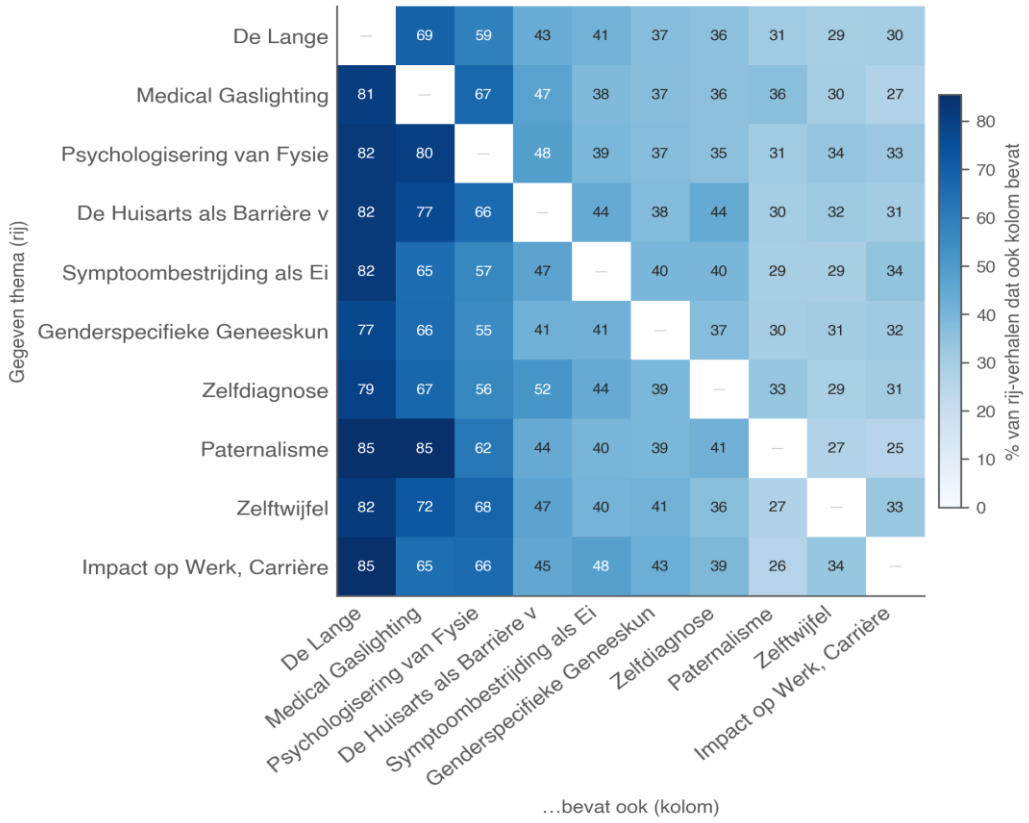
De twintig thema's zijn geclusterd tot **zestien overkoepelende thema's**. De vijf grootste clusters zijn *Psychologisering van fysieke klachten* (406 verhalen, opgebouwd uit 11 onderliggende codes), *Ongeloof en bagatellisering van klachten* (113, 4), *Pleidooi voor gendersensitieve gezondheidszorg* (112, 26), *Langdurige en uitputtende diagnostische zoektocht* (99, 18) en *Gemis van holistische patiëntenzorg* (85, 14).

Hoe thema's samen optreden. Figuur 3 toont de absolute co-occurrence van de tien meest voorkomende inductieve thema's: het aantal verhalen waarin elk paar tegelijk voorkomt. *Niet serieus genomen worden* en *verlies van identiteit* deelden 810 verhalen; *niet serieus genomen worden* en *psychologisering* delen er 778. Deze absolute lezing werd echter gedomineerd door hoe vaak een thema überhaupt voorkwam. Figuur 4 corrigeert hiervoor met een conditionele, asymmetrische lezing: elke cel geeft aan welk percentage van de verhalen met het rij-thema ook het kolom-thema bevat. Daarmee komt de richting van de samenhang in beeld. Negentig procent van de verhalen over decennialange diagnostische vertraging ging ook over *niet serieus genomen worden*; 91% van de verhalen over *patiënt-als-expert* en *zelfadvocacy* doet dat eveneens; 88% van de verhalen over *symptoombestrijding zonder bronaanpak* ook. De omgekeerde percentages waren lager, omdat de grotere thema's ook zonder die specifieke deelervaring voorkwamen. De *diagnostische vertraging*, de *gedwongen zelfadvocacy* en de *symptoombestrijding* waren met andere woorden ingebed in het overkoepelende patroon van miskenning, niet andersom.

Figuur 3 *Absolute co-occurrence van de top-tien inductieve thema's. Celwaarden tonen het aantal verhalen waarin beide thema's aanwezig zijn*

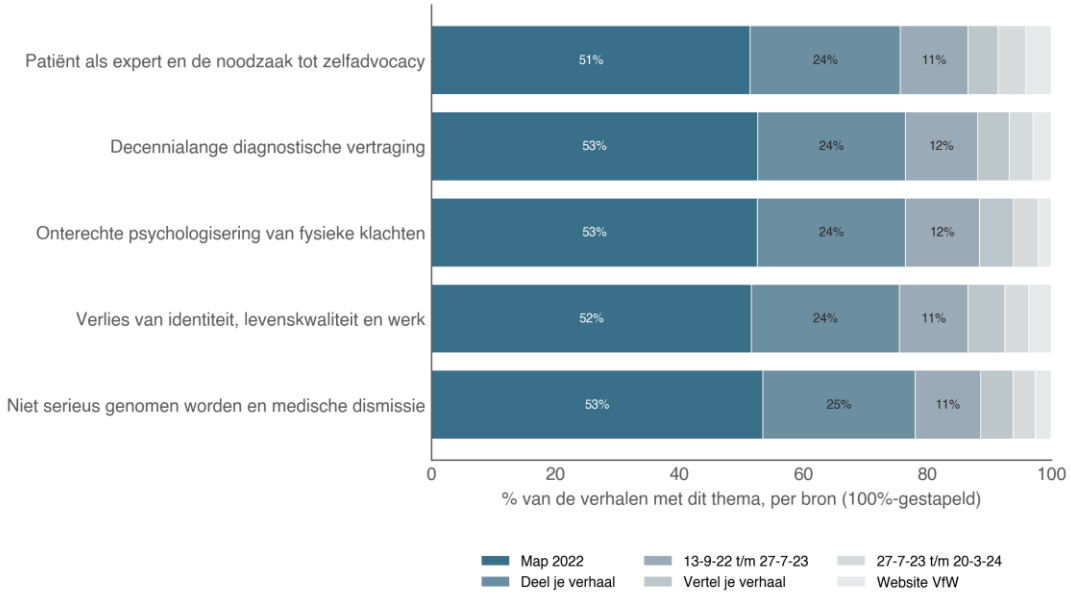


Figuur 4 *Relatieve (conditionele) co-occurrence. Lees per rij: van de verhalen met het rij-thema bevatte dit percentage ook het kolom-thema. De matrix is asymmetrisch en haalt de dominantie van absoluut grote thema's weg*



Bias-controle. Voor de top-vijf thema's is de verdeling over de zes bronnen relatief gemaakt: elke balk sommeerde tot 100% en de segmenten tonen welk aandeel van de verhalen met dat thema uit elke bron afkomstig is. Figuur 5 laat zien dat die verdelingen voor alle top-thema's vrijwel identiek zijn, en sporen met de algemene bronsamenstelling van het corpus. Geen enkel thema was exclusief gebonden aan één bron of disproportioneel oververtegenwoordigd. Leeftijd was voor een groot deel van de verhalen onbekend (de vrije narratieven bevatten dit veld vrijwel nooit); voor de overige verhalen volgden alle top-thema's de algemene leeftijdsverdeling. Het ontbreken van een leeftijdssignaal in de meerderheid van het corpus was een echte beperking voor stratificatie en vraagt om aanvullende metadata bij de bronnen.

Figuur 5 *Verdeling van de top-vijf inductieve thema's over de zes bronbestanden, als percentage binnen elk thema (100%-gestapeld). Dit normaliseerde voor de uiteenlopende grootte van de bronnen en maakte zichtbaar of een thema oververtegenwoordigd was in één bron*



4.2.4 Prominentie, frequentie en belang

In de thematische analyse is hoe vaak een thema voorkomt iets anders dan hoe belangrijk het is. Dat is een bewust uitgangspunt: de gangbare praktijk (Braun & Clarke; de saliency-benadering van Buetow, 2010) waarschuwt ervoor om frequentie gelijk te stellen aan belang. Een thema dat in bijna elk verhaal terloops wordt genoemd hoeft niet de kern van die verhalen te zijn, terwijl een thema dat minder vaak opduikt het verhaal volledig kan beheersen op het moment dat het wel speelt. Om hieraan recht te doen is elk thema langs twee onafhankelijke assen gelegd. De eerste is prevalentie: in hoeveel van de 1.285 verhalen komt het thema voor. De tweede is conditionele centraliteit: van de verhalen waarin het thema voorkomt, in welk aandeel vormt het de centrale klacht (prominentielabel 'centraal'), tegenover een bijkomende of zijdelingse rol. Beide assen zijn berekend voor zowel het deductieve als het inductieve spoor.

Vier klassen. Het kruisen van beide assen levert vier klassen op (Tabel 3). Thema's die vaak voorkomen en meestal centraal staan worden dragend genoemd; zij vormen de ruggengraat van het corpus. Thema's die zelden voorkomen maar centraal zijn wanneer ze opduiken zijn zeldzaam maar kernachtig. Thema's die vaak voorkomen maar zelden het hoofdpunt zijn vormen de alomtegenwoordige context. Thema's die zowel zeldzaam als zijdelings zijn vormen de lange staart.

Tabel 14 *De vier belangklassen (prevalentie x conditionele centraliteit), met voorbeelden uit het deductieve en inductieve spoor. Tussen haakjes: aandeel verhalen waarin het thema aanwezig is/aandeel daarvan waarin het centraal staat*

Klasse	Definitie	Voorbeelden (aanwezig/centraal)
Dragend	Vaak aanwezig en meestal centraal	Meest belastende aspecten (99%/96%); structureel niet serieus genomen worden (87%/65%)
Zeldzaam maar kernachtig	Zelden aanwezig, maar centraal wanneer het speelt	Atypische presentatie en onderdiagnostiek van cardiovasculaire klachten (9%/60%); bagatellisering van (vroege) overgangsklachten (19%/40%)
Alomtegenwoordige context	Vaak aanwezig, maar zelden het hoofdpunt	Informatiebehoefte (48%/4%); lotgenotencontact (28%/1%)
Lange staart	Zelden aanwezig en zijdelings	Diverse zeldzame, zijdelings genoemde thema's (lage prevalentie, lage centraliteit)

Dragende thema's. De dragende thema's bevestigen het beeld uit de vorige paragrafen. Aan de deductieve kant springen de meest belastende aspecten (in 99% van de verhalen aanwezig, en daarvan in 96% centraal) en het gevoel niet gehoord en serieus genomen te worden bij de diagnose (98% aanwezig, 82% centraal) eruit. Aan de inductieve kant zijn dat niet serieus genomen worden (87% aanwezig, 65% centraal) en de onterechte psychologisering van fysieke klachten (53% aanwezig, 52% centraal). Dit zijn de thema's die zowel breed gedragen als zwaarwegend zijn.

Frequentie is niet hetzelfde als belang. Het scherpst wordt dit onderscheid zichtbaar bij twee deductieve vragen met een vrijwel gelijke prevalentie maar een tegengestelde centraliteit. Of een vrouw zich gehoord voelt door de specialist komt in 78% van de verhalen voor en is dan in 55% van die gevallen de kern van het verhaal. De informatiebehoefte komt met 48% ook regelmatig voor, maar is slechts in 4% van die gevallen centraal: ze wordt vaak genoemd, maar bijna altijd als bijzaak. Hetzelfde patroon zien we bij lotgenotencontact (in 28% van de verhalen genoemd, vrijwel nooit centraal) en bij het inductieve thema symptoombestrijding zonder bronaanpak (40% aanwezig, 10% centraal). Deze alomtegenwoordige context kleurt de verhalen, maar drijft ze niet; een ranglijst op frequentie alleen zou ze ten onrechte bovenaan zetten.

Zeldzaam maar kernachtig. Anders dan in de eerste, kleinere analyse is deze vierde en meest intrigerende klasse nu niet langer leeg. Het inductieve thema atypische presentatie en onderdiagnostiek van cardiovasculaire klachten komt in slechts 9% van de verhalen voor, maar staat in 60% van díe verhalen centraal, een klassiek zeldzaam-maar-doorslaggevend patroon dat aansluit bij de bekende onderdiagnostiek van hart- en vaatklachten bij vrouwen. Aan de rand van deze klasse ligt een cluster hormonale en levensfasegebonden thema's met hetzelfde gedrag: de bagatellisering van (vroege) overgangsklachten (19% aanwezig, 40% centraal), de normalisering van invaliderende menstruatiepijn (17% aanwezig, 38% centraal) en zwangerschap en bevalling als trigger voor langdurige klachten (11% aanwezig, 33% centraal). Juist deze vrouwspecifieke ervaringen komen zelden voor maar wegen zwaar wanneer ze spelen, precies de patronen die een ranglijst op frequentie alleen zou missen.

4.3 Beknopte casusbeschrijvingen

In Bijlage B worden vijf willekeurige, geanonimiseerde en representatieve casussen uitgebreid beschreven. Deze verhalen worden hieronder beknopt weergegeven met vetgedrukt de thema's die ook uit de kwalitatieve en de kwantitatieve analyses naar voren zijn gekomen.

Casus 1: vrouw, 20-25 jaar

Zeven jaar lang menstruatie- en vermoeidheidsklachten die door diverse huisartsen werden toegeschreven aan depressie en stress, ook al viel ze flauw van de pijn. Zes verschillende psychofarmaca leidden tot ernstige bijwerkingen onder meer door een te hoge dosering. Kortom: **psychologisering van fysieke klachten** in een vroege levensfase, met iatrogene schade en onderbroken studie- en werkcarrière als gevolg.

Casus 2: vrouw, 40-45 jaar

Meer dan 10 jaar tussen het begin van de klachten en de juiste diagnose vervroegde overgang; kort na de bevalling kreeg zij het label 'bipolaire stoornis' en jarenlang antipsychotica die haar in een zombie veranderden. Meerdere artsen wezen haar eigen verklaring (vervroegde overgang) af met 'je bent te jong', en haar mans observaties werden serieuzer genomen dan haar eigen. Een voorbeeld van decennialange **diagnostische vertraging**, iatrogene schade en het **verlies** van carrière, financiële onafhankelijkheid en tien levensjaren.

Casus 3: vrouw, 26-55 jaar

Meerdere burn-outs, hormonale ontregeling na anticonceptie, ernstige complicaties na de keizersnede; telkens werd de fysieke uitval bestempeld als 'perfectionisme' of stress, en moest zij 'trekken' om SOLK überhaupt als diagnose te krijgen. Geen enkele specialist nam verantwoordelijkheid voor de schade. Toont hoe **gefragmenteerde** zorg en **bagatellisering** iatrogene schade veroorzaken en de patiënt dwingen tot eigen regie.

Casus 4: vrouw, 26-55 jaar

Klachten jarenlang geduid als onverklaarbaar dan wel psychisch voordat in het buitenland endometriose werd vastgesteld evenals onbehandelde trauma's. Zij had zelf via internet de juiste verdenking gevonden, maar zorgverleners wuifden dat weg als 'symptoom van een psychische aandoening'. Illustreert de gedwongen **zelfadvocacy** en de financiële en relationele **schade** van misbehandeling.

Casus 5: vrouw, 26-55 jaar

Meer dan 10 jaar met opiaten en arbeidsongeschiktheid voordat na één grondig onderzoek de diagnose endometriose en adenomyose graad 4 werd gesteld met verklevingen op meerdere plekken. Eerdere artsen hadden de klachten **afgedaan** als PDS, functionele dyspepsie of 'gewoon gezond'.

Deze casusbeschrijvingen zijn illustratief voor de in dit onderzoek gevonden thema's en ervaren knelpunten. De verhalen hebben met elkaar gemeen dat de vrouwen lang rond hebben gelopen met miskende klachten, daar in een aantal gevallen verkeerd voor behandeld zijn en dat ze er op meerdere leefdomeinen langdurige en soms onomkeerbare (iatrogene) schade van ondervonden hebben.

5 Samenvatting en discussie

5.1 Samenvatting

De vrouwen die via de website van Voices for Women de enquête over hun onverklaarde gezondheidsklachten hebben ingevuld rapporteerden een veelheid aan gezondheidsklachten zowel op psychisch als lichamelijk gebied; naast vermoeidheid en pijn werden ook slaapproblemen en stemmingswisselingen veelvuldig als klachten gerapporteerd. De meeste vrouwen hadden deze klachten al 10 jaar of langer en hadden een jarenlange zoektocht doorlopen naar de juiste diagnose en behandeling. Terwijl de diagnosen burn-out/stress, depressie en de overgang in eerste instantie het meest vaak werden gesteld, werden - naast de overgang - ook auto-immuunziekten en endometriose als uiteindelijke, bijgestelde diagnosen genoemd. Deze primaire psychologisering van de gezondheidsklachten kwam ook in de kwalitatieve analyses naar voren.

De enquêtes wezen ook uit dat veruit de meeste respondenten zich niet gehoord voelden bij huisarts noch medisch specialist. Het gevoel niet serieus te worden genomen werd ook bevestigd door de uitkomsten van de kwalitatieve analyse van de ervaringsverhalen; deze verhalen waren overwegend negatief gekleurd. Daarnaast kwamen daarin thema's als verlies van identiteit, kwaliteit van leven en werk; onderschatting van invloed van hormonaal-bepaalde levensfasen; moeizame communicatie; hokjesgeest van medisch specialisten; en decennialange diagnostische vertraging, naar voren. Ook werden in de ervaringsverhalen aspecten genoemd die vrouwen in de gezondheidszorg misten: een holistische blik; dat er naar hen geluisterd wordt; respect voor de expertise van de vrouw zelf; kennis van het vrouwenlichaam; en betrouwbare, gendersensitieve informatie. Deze thema's kwamen grotendeels in alle leeftijdsgroepen naar voren, hoewel de inhoud wel wat verschuift over de levensfasen heen. De resultaten van de kwalitatieve analyses bekrachtigden dus in meerdere opzichten die van de kwantitatieve analyses.

Tenslotte zijn de gezondheidsklachten volgens nagenoeg alle respondenten op de enquête van invloed op meerdere levensdomeinen en dus ook belastend voor hun sociale relaties en financiële situatie. Aanvullend wezen de kwalitatieve analyses op ervaren verlies van ouderrol en werk.

5.2 Discussie

De vrouwen die via de website van Voices for Women hun ervaringen met de gezondheidszorg hebben gedeeld hebben veel gemeen; zij lopen veelal al jaren met hun klachten rond, voelen zich onvoldoende gehoord en serieus genomen door de behandelende zorgverleners en hebben de ervaring dat hun lichamelijke klachten snel als psychisch worden bestempeld en zijzelf als aansteller; een etiket waar veel vrouwen uiteindelijk ook zelf in gaan geloven (*medical gaslighting*). De diagnosen burn-out/stress en depressie worden veelvuldig gesteld, maar lijken ook geregeld te worden bijgesteld nadat vrouwen de weg hebben gevonden naar meer uitgebreid onderzoek. De ervaren gezondheidsklachten zijn voor de respondenten dusdanig ernstig dat ook het sociale leven, relaties en werk van deze vrouwen eronder lijden. Deze bevindingen zijn niet nieuw, maar de combinatie van de enorme impact, de geringe aandacht voor de invloed van hormonale factoren en het gevoel te worden weggezet, die uit de ervaringen naar voren komt, bevestigt andermaal de noodzaak tot meer onderzoek en verspreiding van kennis over het vrouwenlichaam en een meer holistische blik.

De oplossingen die thans uit de literatuur naar voren komen bieden enige houvast voor de richting waarin aandacht voor gezondheidsklachten van vrouwen op zou moeten gaan, maar lijken vooralsnog te summier. Zo is bijvoorbeeld PCOS een van de gynaecologische diagnosen die, naast de overgang en endometriose, na uitgebreid onderzoek geregeld bij vrouwen met langdurige klachten wordt gesteld. Recent is de term PCOS (polycysteusovariumsyndroom) - waarbij de C onterecht verwees naar het bestaan van cystes - vervangen door PMOS (polyendocrien metabool ovariumsyndroom) (Teede et al, 2026). Een van de auteurs van het artikel van Teede et al verwacht dat deze naamswijziging zal leiden tot 'snellere diagnostiek' en 'gerichter onderzoek' naar de langetermijngevolgen (Voorkamp & van Twillert, 2026). Gezien de ervaren bagatellisering van vooral lichamelijke klachten die vrouwen in het huidige onderzoek rapporteren, is het de vraag of deze verwachtingen zullen uitkomen. Diagnosen worden immers pas gesteld nadat klachten als serieus worden erkend en herkend en huisartsen hun patiënten bij een vermoeden op een onderliggende diagnose doorverwijzen voor nader diagnostisch onderzoek en daar lijken vrouwen met langdurige lichamelijke klachten juist in vast te lopen; zij ervaren de huisarts vaak als een hobbel in hun diagnostisch traject. Eerder onderzoek wees ook uit dat, vergeleken met mannen, aan vrouwen vaker labonderzoek wordt aangeboden en minder vaak beeldvormend onderzoek (Ballering et al, 2021). Het belang van het beter oog en oor hebben voor de klachten en ervaringen van vrouwen wordt onderschreven in een recent artikel waarin een neuroloog oproept om vrouwen met hoofdpijnklachten niet af te schepen met paracetamol of 'het vermijden van triggers', maar oog te hebben voor de invloed van hormonale factoren en vooral te luisteren naar vrouwen ter preventie van onderdiagnostiek van - in dit voorbeeld - migraine (Sulter, 2026). Een advies dat - mits het navolging krijgt - bij veel vrouwen waarschijnlijk sneller tot een juiste diagnose leidt dan het veranderen van de naam van een diagnose.

Voor het tijdig stellen van een (juiste) diagnose, is de kwaliteit van de zorgverlener-patiënt communicatie cruciaal. Onderzoek naar communicatie in de spreekkamer van huisarts en medisch specialist wijst uit dat vrouwelijke patiënten lange tijd anders bejegend werden dan mannelijke patiënten, hoewel deze verschillen de laatste jaren wel lijken af te nemen (Noordman & van Dulmen, 2016; Plug et al, 2021). Zo blijkt bijvoorbeeld dat vrouwen in medische consulten meer de verbinding zoeken met de ander, onder andere door het gebruik van *tag-questions*, waarmee ze rechtvaardiging voor hun gedachten en ideeën zoeken (van Dulmen & Plug et al, 2022) maar dat zij ook vaker geïnterrumpeerd worden (Plug et al, 2022). Een kritische reflectie op de eigen wijze van communiceren zowel aan de kant van de zorgverlener als aan de kant van de patiënt, kan helpen om meer bewustzijn te creëren over het belang en de impact van communicatief gedrag en om de wijze van communiceren te verbeteren. Meer onderzoek naar hoe juist vrouwen geholpen kunnen worden om op een effectieve en assertieve manier hun klachten en ervaringen over te brengen, is daarvoor van belang. Want - zo wijst recent onderzoek uit - assertiviteit doet ertoe, ook voor het verkrijgen van een verwijzing (Cariot et al, 2025).

5.2.1 Waarde van kwalitatieve tweesporenaanpak op basis van AI

In dit onderzoek is gebruik gemaakt van AI voor het identificeren van onderliggende en gemeenschappelijke thema's in de ervaringsverhalen van vrouwen met langdurige gezondheidsklachten. Deze aanpak werd zowel op een deductieve als op een inductieve manier vormgegeven. De uitkomsten van deze tweesporenaanpak zijn vervolgens met elkaar vergeleken. Dit leverde twee interessante observaties op:

Convergentie rond een herkenbare structuur. De vaakst gescoorde deductieve vragen (gehoord worden bij de diagnose, communicatie met zorgverleners, behoeften en verbeterpunten) vonden

hun inductieve tegenhanger in de drie grootste thema's: niet serieus genomen worden, onterechte psychologisering en de patiënt als gedwongen expert. Hetzelfde sub-cluster, psychologisering, dook bovendien op binnen vrijwel elke deductieve vraag waar zij relevant was: van de communicatie met zorgverleners (n=22) en de huisartsen-ervaring (n=30) tot de samenhang lichaam-psyche (n=35) en de meest belastende aspecten (n=18). Dat versterkte het vertrouwen dat het codeboek de centrale ervaringspatronen vangt.

Divergentie: wat het inductieve spoor toevoegt. Vier inductieve thema's markeren een maatschappelijk-economische as die in de oorspronkelijke vragenlijst onderbelicht bleef: *Verlies van identiteit, levenskwaliteit en werk* (n=934), *Patiënt als expert en de noodzaak tot zelfadvocacy* (n=743), *Sociaal isolement en maatschappelijk onbegrip* (n=304) en *Strijd met betrokken instanties* (n=76). Voor toekomstig onderzoek is dit relevant; een eventuele uitbreiding van de registry zou deze constructen aan de impact-categorie en aan de patiëntervaring-categorie kunnen toevoegen, zodat zij ook gestructureerd worden gescoord.

Concluderend kan gesteld worden dat de uitkomsten van het deductieve spoor worden bekrachtigd door die van het inductieve spoor en vice versa. Dit versterkt de betrouwbaarheid van de uitkomsten. In vervolgonderzoek wordt een dergelijke tweesporenaanpak dan ook aanbevolen.

5.2.2 Input voor VWS Werkagenda Vrouwengezondheid

VWS heeft voor de Werkagenda Vrouwengezondheid drie actielijnen uitgezet:

1. **Herkennen en erkennen van klachten**
   Deze actielijn richt zich op het vergroten van kennis en bewustzijn over aandoeningen of klachten die alleen, voornamelijk of anders bij vrouwen voorkomen, inclusief de invloed van de vrouwelijke hormonale dynamiek op het ontstaan, verloop of behandeling van ziektes. Daarnaast richt deze actielijn zich op het creëren van meer aandacht voor de invloed van sociaal-maatschappelijke factoren, zoals de combinatie van betaald en onbetaald werk, sociale normen, bestaanszekerheid, en de impact hiervan op de gezondheid van vrouwen. Een belangrijk onderdeel is het beter laten landen van kennis die er al is, zodat vrouwen en meisjes serieus worden genomen, zich gehoord voelen en sneller de juiste diagnose én passende zorg krijgen.
2. **Verbeterde kennisontwikkeling, implementatie en innovatie**
   Deze actielijn omvat onder andere het stimuleren van kennisontwikkeling en zorgen dat zowel bestaande als nieuwe kennis terechtkomen op de juiste plekken, zoals in richtlijnen, behandelprotocollen, (preventie)beleid en onderwijs. Daarnaast richt deze actielijn zich op het vergroten en gerichter inzetten van financiering voor aandoeningen en problemen die alleen of voornamelijk bij vrouwen voorkomen, en op het structureel verankeren van man/vrouw-verschillen in gezondheid, zorg en welzijn, bijvoorbeeld in criteria voor onderzoeksfinanciering en bij de ontwikkeling en implementatie van interventies en innovaties.
3. **Verbeterde infrastructuur en uitwisseling**
   Deze actielijn richt zich op het verbeteren van de infrastructuur en interdisciplinaire uitwisseling rondom de gezondheid van vrouwen. Er is grote behoefte aan een bredere, holistische blik op vrouwengezondheid, waarbij niet alleen medische, maar ook sociale en maatschappelijke factoren worden meegenomen. Om dit te realiseren, is meer uitwisseling en samenwerking vereist tussen de verschillende (medische en sociale) disciplines, om te voorkomen dat er in silo's wordt gewerkt. Betere structuren voor het delen van data, het opnemen van parameters rondom vrouwengezondheid en het delen van best practices vormen hierin een belangrijk onderdeel.

Het onderhavige onderzoek biedt primair aanknopingspunten voor actielijn 1: Het herkennen en erkennen van klachten. Slechts één op de vier vrouwen die hun ervaringen hebben gedeeld door

middel van het invullen van de enquête geeft in dit onderzoek aan zich gehoord te hebben gevoeld bij de zorgverlener, het merendeel heeft negatieve ervaringen en voelt zich weggezet en hun klachten gebagatelliseerd. In alle ervaringsverhalen komen thema's als niet serieus genomen worden en het gevoel dat er niet naar vrouwen geluisterd wordt ook naar voren. Deze uitkomsten en die van meer uitgebreid vervolgonderzoek kunnen gebruikt worden om zorgverleners meer bewust te maken van deze ervaringen. Daarnaast is meer inzicht in wat zorgverleners en vrouwen zelf kan helpen om de zorg te verbeteren nodig om deze actielijn verder te voeden. Daarvoor is input vanuit actielijn 2 cruciaal. Voor actielijn 3 biedt dit onderzoek onderbouwing omdat het onderzoek laat zien dat een meer holistische aanpak lijkt te ontbreken en lichamelijke klachten structureel gepsychologiseerd worden en de grote maatschappelijk-economische gevolgen thans onderbelicht blijven.

5.2.3 Methodologische kanttekeningen

Ondanks de rijkheid aan data was het in het korte bestek van het project niet mogelijk om meer diepgaande analyses te verrichten. Er zijn nog wel veel mogelijkheden voor verder onderzoek. Zo zou het interessant kunnen zijn om te onderzoeken welke stappen vrouwen door de tijd heen hebben doorlopen alvorens bij er bij hen een juiste diagnose gesteld werd. Het onderzoeken van de tijdsduur tot het stellen van de juiste diagnose zou daarbij meegenomen kunnen worden. Maar ook de samenhang van de (lichamelijke en psychische) klachten alsmede de maatschappelijke impact van klachten en diagnosen zijn onderwerpen die om nader onderzoek vragen. De enquêtes waren niet ontworpen voor wetenschappelijk onderzoek, hadden verschillende vragen en maakten geen gebruikt van gevalideerde schalen. Veel vragen waren bovendien niet eenduidig geformuleerd. Zo was bij de vraag naar de gestelde diagnose niet duidelijk wie die diagnose gesteld heeft en bij de vraag naar de *juiste* diagnose, niet duidelijk volgens wie die juist zou moeten zijn. Gezien de beperkingen van de enquêtes, konden niet alle onderzoeksvragen beantwoord worden. Door de extreem korte looptijd van het project was het nu niet mogelijk om de antwoorden op de open vragen inhoudelijk te analyseren.
Bovendien is de zelfselectie van het corpus door de aard van het meldpunt onvermijdelijk en kleurt de gevonden patronen; hoe representatief de ervaringen zijn voor alle Nederlandse vrouwen met onverklaarde gezondheidsklachten is dus niet te zeggen. Ook is leeftijd voor 55% van het corpus onbekend, wat stratificatie op die dimensie beperkt; aanvullende metadata bij de bronnen zou dit deels kunnen oplossen. Tenslotte scoort de registry-vraag over interacties van vrouwen van kleur over het hele corpus minimaal (1,2%). Uit de verhalen zijn meestal geen demografische karakteristieken af te leiden, dus we kunnen niet met zekerheid zeggen of dit een steekproef-eigenschap reflecteert of met name te maken heeft met de manier van schrijven.

In een uitgebreider vervolgproject zou de inhoud van de enquête daarom opnieuw tegen het licht gehouden moeten worden. Ook zouden meer achtergrondkenmerken van de respondenten uitgevraagd moeten worden.

6 Literatuur

Ballering AV, Muijres D, Uijen AA, Rosmalen JGM, Olde Hartman TC. Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: an observational cohort study J Psychosom Res 2021 Oct;149:110589

Ballering AV, Rosmalen JGM, Olde Hartman TC. Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. J Am Board Fam Med 2022 Jan-Feb;35(1):73-84

Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Sage, 1998

Braun V, Clarke V. Thematic Analysis: A Practical Guide. Sage, 2022

Cariot L, Noordman J, Leemrijse C, Ginkel M van, Dulmen S van. Patient assertiveness and visit outcome in primary care: An observational study. PEC 2025 April 17;137:108789

Calissendorff J, Cramon PK, Hallengren B, Khamisi S, Lantz M, Planck T, Sjölin G, Wallin G, Holmberg M. Long-Term Outcome of Graves' Disease: A Gender Perspective. Womens Health Rep (New Rochelle) 2023 Oct 9;4(1):487-496

Dulmen S van, Plug I. Mannelijke en vrouwelijke huisartsen communiceren anders, toch? In: Gendersensitieve huisartsgeneeskunde; een handboek voor de praktijk (T.Lagro-Janssen & D.Teunissen, Eds). BSL 2022

Eggermont D, Smit MAM, Kwestroo GA, Verheij RA, Hek K, Kunst AE. The influence of gender concordance between general practitioner and patient on antibiotic prescribing for sore throat symptoms: a retrospective study. BMC Fam Pract 2018 Nov 17;19(1):175

Eggermont D, Kunst AE, Hek K, Verheij RA. Gender and age concordance between patient and GP: an observational study on associations with referral behaviour. BJGP Open 2022 Dec 20;6(4): BJGPO.2022.0091

Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology 2013; 13: 117. https://doi.org/10.1186/1471-2288-13-117

Harrilal-Maharaj S. Diagnostic gender bias and psychologisation in women's healthcare'. *Pre-print, in literatuurreview methodologie*, 2025

Hill, [first author]. 'Three-tier error taxonomy for blinded LLM-assisted thematic coding in healthcare'. *Pre-print*, 2026.

Houwen J, Lucassen PL, Stappers HW, Assendelft WJ, van Dulmen S, Olde Hartman TC. Improving GP communication in consultations on medically unexplained symptoms: a qualitative interview study with patients in primary care. Br J Gen Pract 2017 Oct;67(663):e716-e23

Kidd IJ, Carel H. Epistemic Injustice and Illness. In *The Routledge Handbook of Epistemic Injustice*. IJ Kidd, J Medina, G Pohlhaus (Red). Routledge, 2017

National Institute for Health and Care Research. *Feeling unheard: women's experiences of healthcare*. NIHR Evidence Collection. https://evidence.nihr.ac.uk, 2024

Noordman J, Dulmen S van. The consequences of task delegation for the process of care: female patients seem to benefit more. Women Health 2016;56(2):194-207

Olde Hartman TC, Rijswijk E van, Dulmen S van, Weel-Baumgarten E van, Schellevis F, Lucassen PLBJ, Weel C van. How general practitioners talk with patients with persistent medically unexplained symptoms. A qualitative study of video-recorded consultations. Patient Educ Couns 2013; 90: 354-360

Plug I, Stommel W, Lucassen PLBJ, olde Hartman TC, van Dulmen S, Das E. Do Women and Men Use Language Differently in Spoken Face-to-Face Interaction? A Scoping Review. Review of Communication Research 2021; 9: 43-79

Plug I, Dulmen S van, Stommel W, Olde Hartman TC, Das E. When physicians' interruptions do not harm. An empirical quantitative analysis of interruptions in clinical practice interactions. Ann Fam Med 2022; 20: 423-429

Stortenbeker I, Stommel W, Dulmen S van, Lucassen P, Das E, olde Hartman T. Linguistic and interactional aspects that characterize consultations about medically unexplained symptoms: A systematic review. J Psychosom Res 2020; 132: 109994

Sulter G. Migraine verdient meer dan schouderophalen. Medisch Contact 4 juni 2026; 23: 25-26

Teede HJ, Khomami MB, Morman R, Laven JSE, Joham AE, Costello MF, Patil M, Rees DA, Berry L, Cree MG, Zhao H, Norman RJ, Dokras A, Piltonen T; Global Name Change Consortium. Polyendocrine metabolic ovarian syndrome, the new name for polycystic ovary syndrome: a multistep global consensus process. Lancet 2026 Jun 6;407(10545):2329-2339

Voorkamp V, van Twillert M. PCOS: hoe de houdbaarheidsdatum van een ziektenaam kan verstrijken. Medisch Contact 4 juni 2026; 23: 10-11

Bijlage A Registry met vragen per type analyse

Categorie	Onderzoeksvraag (kwalitatief hoofdvraag)	Type analyse
1. Beschrijving onderzoekspopulatie	Welke demografische en sociaal-culturele kenmerken beschrijven de onderzoekspopulatie?	**Kwantitatief**
	Hoe is de onderzoekspopulatie verdeeld over de verschillende leeftijdsgroepen?	**Kwantitatief**
	Wat is het aandeel vrouwen van kleur binnen de onderzoekspopulatie, en hoe is dit verdeeld over de leeftijdsgroepen?	**Kwantitatief**
	In welke landen hebben de gezondheidsklachten zich voorgedaan en hoe is de onderzoekspopulatie over deze landen verdeeld?	**Kwantitatief**
2. Aard en duur van klachten	Welke lichamelijke klachten komen naar voren in de ervaringen van vrouwen?	**Kwantitatief**
	Welke psychische klachten komen naar voren, en hoe ervaren vrouwen dat deze zich verhouden tot lichamelijke klachten?	**Kwantitatief**
	Hoe wordt de samenhang tussen lichamelijke en psychische klachten ervaren door vrouwen én geïnterpreteerd door behandelaars?	**Mixed**
	Hoe beschrijven vrouwen het verloop van hun klachten en de periode tot het zoeken van hulp of verkrijgen van een diagnose?	**Kwantitatief**
	Welke patronen in klachten komen naar voren in relatie tot vrouwelijke levensfasen en hormonale veranderingen, en hoe ervaren vrouwen deze?	**Mixed**
	Welke patronen zijn zichtbaar in de aard van klachten per leeftijdsgroep (10–25, 26–55, 55+)?	**Kwantitatief**
3. Diagnoseroute en erkenning	Welk aandeel van de vrouwen heeft (nog) geen diagnose gekregen, en welk aandeel heeft een onjuiste eerste diagnose gehad?	**Kwantitatief**
	Welke diagnoses worden het vaakst gesteld, en welke diagnoses blijken achteraf onjuist?	**Kwantitatief**
	Hoelang duurt het gemiddeld tussen eerste klachten en een juiste diagnose, en verschilt dit per leeftijdsgroep of etnische achtergrond?	**Kwantitatief**
	Hoe ervaren vrouwen het proces van diagnostiek en de rol van zorgverleners daarin? Voelen zij zich gehoord, serieus genomen en begeleid?	**Kwalitatief**
	Welke rol spelen psychische labels (burn-out, depressie, angststoornis) in het traject naar een lichamelijke diagnose?	**Mixed**

Categorie	Onderzoeksvraag (kwalitatief hoofdvraag)	Type analyse
4. Zorggebruik en zorgprofessionals	Hoe vaak bezoeken vrouwen de huisarts met hun klachten, en hoe beschrijven vrouwen hun contactmomenten met de huisarts en de ervaren kwaliteit daarvan?	**Mixed**
	Welke specialisten worden het vaakst bezocht, en hoeveel bezoeken zijn er gemiddeld nodig per specialist?	**Kwantitatief**
	In hoeverre voelen vrouwen zich gehoord door specialisten, en verschilt dit van de ervaring bij de huisarts?	**Mixed**
	Hoe beschrijven vrouwen de communicatie met zorgverleners, en welke elementen dragen bij aan positieve of negatieve ervaringen?	**Kwalitatief**
	Welke signalen in de verhalen van vrouwen kunnen huisartsen helpen om klachten eerder te herkennen als vrouwspecifiek?	**Kwalitatief**
	Hoe ervaren vrouwen van kleur hun interacties met zorgverleners, en welke verschillen zien zij ten opzichte van andere vrouwen?	**Mixed**
5. Behandeling en medicatie	Hoeveel vrouwen rapporteren onnodige operaties, behandelingen of medicatie te hebben ondergaan?	**Kwantitatief**
	Welke medicatie wordt het vaakst als 'onnodig' bestempeld, en welke patronen zijn hierin zichtbaar?	**Kwantitatief**
	Hoe beschrijven vrouwen de mate waarin rekening wordt gehouden met hun hormoonhuishouding bij het voorschrijven van medicatie?	**Mixed**
	Hoe ervaren vrouwen de wijze waarop wordt omgegaan met hun meldingen van bijwerkingen, en hoe beschrijven zij hun besluitvorming rondom het stoppen met medicatie? En welk aandeel van de vrouwen geeft aan te zijn gestopt met medicatie vanwege bijwerkingen?	**Mixed**
6. Maatschappelijke en persoonlijke impact	In welke mate is er sprake van arbeidsuitval door onverklaarde gezondheidsklachten (deels en volledig), en wat is de gemiddelde duur hiervan?	**Kwantitatief**
	Op welke manieren beïnvloeden klachten fr verschillende levensdomeinen (gezin, sociale contacten, financiën) van vrouwen?	**Mixed**
	Wat beschrijven vrouwen als de meest belastende aspecten van hun zorg- en ziekteproces?	**Kwalitatief**
	Hoe beïnvloeden de gezondheidsklachten het sociale leven en de relaties van vrouwen?	**Kwalitatief**

Categorie	Onderzoeksvraag (kwalitatief hoofdvraag)	Type analyse
7. Behoeften en verbeterpunten	Wat zijn de meest genoemde verbeterpunten in de zorg voor vrouwen met onverklaarde klachten?	**Kwalitatief**
	Welke behoeften noemen vrouwen als het gaat om communicatie, erkenning en begeleiding door zorgverleners?	**Kwalitatief**
	Welke informatiebehoeften hebben vrouwen (bijv. over diagnoses, hormoonwerking, behandelopties)?	**Kwalitatief**
	Welke rol spelen lotgenotencontact en zelfhulp in het omgaan met onverklaarde klachten?	**Kwalitatief**
8. Overkoepelende patronen	Welke verschillen en overeenkomsten in klachtbeleving, zorggebruik of uitkomsten komen naar voren in ervaringen tussen vrouwen van kleur en andere vrouwen?	**Mixed**
	Welke drie tot vijf overkoepelende knelpunten komen het sterkst naar voren uit zowel kwantitatieve als kwalitatieve data?	**Mixed**
9. Verzoek ZonMw	In welke mate sluiten de gevonden knelpunten aan bij de drie actielijnen van VWS (herkennen & erkennen; kennisontwikkeling; infrastructuur & samenwerking)?	**Kwalitatief**
	Onderbouw de analyse met maximaal zes korte, geanonimiseerde patiëntverhalen die spreiding tonen over leeftijdsgroepen, domeinen en achtergronden.	**Kwalitatief**

LEGENDA	
Kwalitatief	Thematische analyse op open tekstvelden
Kwantitatief	Beschrijvende statistiek op gesloten enquêtevragen
Mixed	Combinatie van kwantitatieve en kwalitatieve analyse

Bijlage B Leren over LLM-beoordeling in thematische analyse

Het handmatig naast elkaar leggen van model- en mensoordeel maakte zichtbaar hoe lastig het is om grote, complexe en meerduidige ervaringsverhalen te coderen, voor de menselijke beoordelaar evenzeer als voor het taalmodel. Drie patronen kwamen terug. Ten eerste ambiguïteit van bepaalde woordkeuze. Zo maakt de ‘de arts’ of ‘de dokter’ zelden expliciet of het om de huisarts of een specialist gaat, wat de vraagtoewijzing vertroebelt; in de laatste versie is dat met scope-notes aangescherpt, zichtbaar in de verschoven huisarts- en specialistpercentages. Ten tweede zagen we soms conceptuele overlap: een uiting als ‘ik wil gehoord worden’ raakt meerdere registryvragen tegelijk, waarbij het niet altijd duidelijk was of het op alle vragen centraal stond of dat het bij een vraag eerder neutraal gekleurd was of toch negatief. Ten derde impliciet versus extrapolatie: mensen en taalmodellen lezen ook tussen de regels door. Veel van de vragen worden niet expliciet beantwoord of niet expliciet als negatief bestempeld, maar bij het lezen van de hele tekst kan je er toch duiding aan geven. Het viel ons op dat deze duiding subjectief en multi-interpretabel blijft. Voorbeeld: zowel model als mens neigt ertoe een psychische last of informatiebehoefte af te leiden waar de tekst die niet expliciet benoemt. “Niet gehoord worden” zal vaak ook psychisch belastend zijn, ondanks dat dit niet expliciet werd genoemd. Bij de menselijke validatie werd het duidelijk dat coderen uitdagend is bij zulke rijke verhalen. Daarbij was de menselijke codering erg belangrijk als diagnostisch instrument om de LLM-logica en prompts aan te scherpen. Tegelijkertijd merkte we bij nadere lezing soms ook dat het model het bij het rechte eind had en dat de mens te streng of letterlijk codeerde. Juist door twee menselijk codeurs dezelfde teksten te laten lezen, werd meer inzicht verkregen in consistentie. Kwantitatief kwam de interbeoordelaarsbetrouwbaarheid op themapresentie uit op een substantiële Cohen’s κ van 0,76 (159 gedeelde cellen), waarbij het model met elke codeur ongeveer even goed overeenkwam als de codeurs onderling (κ = 0,67 en 0,74). Het sentimentoordeel kent een zeer hoge ruwe overeenstemming (85–96%) bij een lagere κ, een kappa-paradox die past bij het overwegend negatieve sentiment in dit corpus.