



Universiteit
Leiden
The Netherlands

Identify, appraise and individualize: clinical practice and prediction models in recurrent pregnancy loss

Youssef, A.

Citation

Youssef, A. (2023, October 10). *Identify, appraise and individualize: clinical practice and prediction models in recurrent pregnancy loss*. Retrieved from <https://hdl.handle.net/1887/3643184>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3643184>

Note: To cite this publication please use the final published version (if applicable).



NEDERLANDSE SAMENVATTING & APPENDICES

NEDERLANDSE SAMENVATTING

Het krijgen van een miskraam is een van de meest voorkomende zwangerschapscomplicaties. Herhaalde miskramen worden doorgaans gedefinieerd als het spontane verlies van twee of meer zwangerschappen voordat de termijn van levensvatbaarheid wordt bereikt (24 weken). Zowel maternale als paternale aandoeningen en risicofactoren zijn geassocieerd met herhaalde miskramen, waaronder het antifosfolipidensyndroom, uterusanomalieën, abnormale karyotypering, sperma-DNA-fragmentatie en leefstijlfactoren. Het exacte aantal koppels met herhaalde miskramen is onbekend.

Ondanks het feit dat herhaalde miskramen al sinds de jaren '30 worden onderzocht, blijven er nog veel vragen onbeantwoord. Het onderliggende mechanisme van herhaalde miskramen is nog steeds niet ontrafeld. Onderzoek gericht op foetale chromosomen van miskraamweefsel toont aan dat in meer dan de helft van de miskramen sprake is van aneuploidie (een verkeerd aantal chromosomen). Dit lijkt geassocieerd te zijn met oudere leeftijd, roken en obesitas, hoewel koppels met een hoog aantal miskramen juist vaker euploïde miskramen krijgen (foetus met het juiste aantal chromosomen). Ondanks uitgebreid onderzoek wordt er slechts in 25% van de koppels met herhaalde miskramen een oorzaak gevonden. Hierdoor blijft het voor medici moeilijk om te verklaren waarom sommige zwangerschappen succesvol verlopen en andere tot een miskraam leiden.

Er is veel discussie rondom het definiëren, onderzoeken en behandelen van herhaalde miskramen, hoewel er verschillende richtlijnen bestaan, waaronder die van de European Society for Human Reproduction and Embryology (ESHRE). Dit geldt des te meer voor koppels met onverklaarde herhaalde miskramen, waarbij geen onderliggende aandoening lijkt te zijn en (experimentele) behandelingen niet worden aanbevolen. Deze onenigheid leidt vaak tot onderzoeken en behandelingen die niet wetenschappelijk onderbouwd zijn, en tot verschillen in de aanpak van herhaalde miskramen tussen verschillende centra.

Deze praktijkvariatie komt tot uiting bij koppels die buiten Nederland zorg zoeken, waar ze uitgebreidere diagnostiek en mogelijke behandelingen kunnen krijgen, hoewel deze vaak wetenschappelijk niet bewezen effectief zijn. Behandelingen die niet bewezen effectief zijn, gaan daarom in tegen

het principe van "niet schaden". Het is logisch dat de psychologische belasting voor stellen met herhaalde miskramen toeneemt na elk zwangerschapsverlies. Het is daarom belangrijk om voldoende zorg te bieden.

Eén van de vormen van ondersteunende zorg bestaat uit het bieden van een toekomstperspectief betreffende de kans op een succesvolle zwangerschap. Deze informatie kan door stellen worden gebruikt bij het nemen van beslissingen over eventuele toekomstige zwangerschapspogingen. In de dagelijkse praktijk van zorg bij herhaalde miskramen bestaan er twee predictiemodellen die worden gebruikt om de kans op toekomstige zwangerschappen te voorspellen, namelijk het model van Lund et al. en het model van Brigham et al. Deze modellen worden genoemd in verschillende internationale richtlijnen. In deze modellen wordt de kans op een succesvolle zwangerschap geschat op basis van leeftijd en aantal eerdere miskramen. Hoewel deze modellen als van hoge methodologische kwaliteit worden beschouwd, hebben ze niet de aanbevolen richtlijnen gevolgd voor het ontwikkelen van een predictiemodel.

Predictiemodellen bieden klinici handvatten om informatie te verstrekken over het optreden van bepaalde uitkomsten. Vier elementen maken predictiemodellen effectief in de dagelijkse praktijk: ontwikkeling, validatie, evaluatie van de impact en updates.

Gewoonlijk zijn predictiemodellen afgeleid van een multivariabel regressiemodel voor dichotome uitkomsten. In de ontwikkelingsfase is het belangrijk dat de database waarop het model is gebaseerd groot genoeg is om betrouwbaarheid te garanderen. In het verleden werd de vuistregel "10 gebeurtenissen per prognostische factor" gebruikt. Er is veel discussie geweest in de wetenschappelijke literatuur over deze regel, waarbij de hypothese is dat het aantal gebeurtenissen per prognostische factor contextspecifiek is en afhankelijk is van meerdere factoren. Daarom heeft Riley et al. een vierstappenprocedure ontwikkeld om een voldoende grote steekproefomvang te berekenen en een model met weinig foutgevoeligheid te waarborgen.

Validatie van een predictiemodel is belangrijk om de generaliseerbaarheid van het model te testen. Hierbij wordt het ontwikkelde model gebruikt in een nieuwe externe dataset, waarbij de voorspellende prestaties van het

model worden onderzocht. Vervolgens kan het model geïmplementeerd worden en kan worden onderzocht hoe het model de patiëntuitkomsten en kosteneffectiviteit beïnvloedt. Tot slot kan het model worden bijgewerkt op basis van nieuwe inzichten.

In een poging om de kwaliteit van zorg voor stellen met herhaalde miskramen te verbeteren, richt dit proefschrift zich op twee hoofdgebieden. In deel I staat de aanwezigheid van verschillende richtlijnen en hun gebruik in de klinische praktijk centraal bij het begrijpen van de zorg voor herhaalde miskramen zoals die vandaag de dag wordt verleend. Het is belangrijk om (inter)nationale verschillen in de zorg voor herhaalde miskramen te bestuderen, omdat praktijkvariatie kan leiden tot een toenemend aantal patiënten dat zorg zoekt bij verschillende (inter)nationale centra, om uitgebreidere onderzoeken en verschillende niet-wetenschappelijk bewezen behandelingen te ondergaan. Deel II van dit proefschrift richt zich op het gebruik van prognostische modellen ter ondersteuning van stellen met herhaalde miskramen. Aangezien begeleiding bij toekomstige zwangerschappen een sleutelrol speelt in de zorg voor herhaalde miskramen, is het belangrijk dat klinici accurate informatie kunnen verstrekken. Na zorgvuldige evaluatie van beschikbare prognostische modellen zal een nieuw prognostisch instrument worden geïntroduceerd.

DIT PROEFSCHRIFT

Hoofdstuk 2 toont een vergelijking van zeven lokale Nederlandse protocollen voor herhaalde miskramen, die zijn vergeleken met de Nederlandse NVOG-richtlijn. Hoewel de gevonden verschillen tussen de protocollen voornamelijk betrekking hadden op de details van het uitvoeren van onderzoeken, weerspiegelt dit het feit dat praktijkvariatie zelfs op lokaal en nationaal niveau aanwezig is.

In **hoofdstuk 3** hebben we praktijkvariatie op internationaal niveau geanalyseerd. Ons doel was om de aanbevolen definities, risicofactoren, onderzoeken en therapieën van drie prominente richtlijnen voor herhaalde miskramen samen te vatten en te vergelijken. We hebben discrepanties gevonden tussen de richtlijnen op al deze genoemde gebieden. Risicofactoren en onderzoeken zijn over het algemeen vergelijkbaar tussen de vergeleken richtlijnen, en de gevonden verschillen kunnen deels worden verklaard door hun ontwikkelingsmethoden en het tijdstip van publicatie,

evenals het gebrek aan sterk bewijs voor sommige klinische aspecten van herhaalde miskramen.

Stellen met herhaalde miskramen zijn zich zeer bewust van onderzoeken en behandelingen die in andere landen worden uitgevoerd, en ze zoeken deze vaak als laatste redmiddel, omdat ze geen succesvolle zwangerschap kunnen bereiken. Er is behoefte aan een universele, op bewijs gebaseerde richtlijn voor herhaalde miskramen om ervoor te zorgen dat stellen met herhaalde miskramen over de hele wereld vergelijkbare en op bewijs gebaseerde onderzoeken en behandelingsopties krijgen. Natuurlijk zijn niet alle aanbevelingen van toepassing op alle bevolkingsgroepen wereldwijd, maar als landen vergelijkbaar zijn op het gebied van medische diensten en bevolkingssamenstelling, kunnen richtlijnen worden samengevoegd en aangepast aan lokale gezondheidszorgstructuren en -organisaties. Een van de belangrijkste hindernissen bij de universele toepassing van één richtlijn is de inconsistentie van de definitiecriteria voor herhaalde miskramen, die lijkt te worden bepaald door meningen in plaats van bewijs. Als dit niet internationaal consistent is, zal dit inherent leiden tot verschillen in de behandeling van herhaalde miskramen. De definitie heeft ook consequenties voor de beschikbare middelen, omdat dit bepaalt wanneer onderzoeken worden ingezet.

De voorgaande hoofdstukken hebben aangetoond dat richtlijnen verschillen, zowel op nationaal als internationaal niveau. Aangezien de ESHRE-richtlijn wordt aangepast voor gebruik in Nederland, hebben we geprobeerd mogelijke belemmeringen bij de implementatie van een nieuwe op bewijs gebaseerde richtlijn te achterhalen. **Hoofdstuk 4** beschrijft een vragenlijstonderzoek dat is uitgevoerd in alle gynaecologiepraktijken in Nederland, waarbij de huidige zorg voor herhaalde miskramen en de visie van klinici daarop zijn geïdentificeerd. We hebben waargenomen dat Nederlandse klinici over het algemeen de op bewijs gebaseerde onderzoeken en therapeutische interventies in de zorg voor herhaalde miskramen naleven, maar dat er ook ruimte is voor verbetering. De belangrijkste verschillen in richtlijnaanbeveling en de praktijk van klinici op het gebied van herhaalde miskramen liggen in de uitgevoerde onderzoeken en behandelingen, zoals karyotypering, screening op erfelijke trombofilie, schildklierfunctie en auto-immuniteit, en β 2-glycoproteïne-antilichaamtesten in de context van het antifosfolipidensyndroom. In dit

hoofdstuk waren we echter niet in staat om de reden achter de verschillen in de zorg voor herhaalde miskramen te onderzoeken. Toekomstige interviewstudies zouden kunnen worden uitgevoerd om te verduidelijken waarom klinici ervoor kiezen om niet op bewijs gebaseerde zorg te bieden.

Hoofdstuk 5 had als doel een implementatiestrategie te testen om de naleving van de richtlijn voor herhaalde miskramen door zorgprofessionals te verbeteren. We hebben vastgesteld dat vier elementen direct verband hielden met een betere naleving: een gespecificeerd medisch dossier, een vragenlijst voor patiënten, een zakkaart en een elektronisch beslissingsprogramma. De naleving was aanzienlijk hoger voor de meeste indicatoren met betrekking tot diagnostiek en counseling. Deze strategieën resulteerden in een kostenverlaging van 206.916 euro per jaar in de vier deelnemende centra. Deze kostenverlaging zou mogelijk nog hoger zijn wanneer dit wordt geëxtrapoleerd naar alle centra die herhaalde miskramen zorg bieden. Uit de literatuur is gebleken dat het verspreiden van richtlijnen alleen onvoldoende is voor een goede naleving. Er bestaan echter geen gouden standaarden voor het ontwikkelen van strategieën ter ondersteuning van de implementatie van bijgewerkte of nieuwe richtlijnen. Overkoepelende organisaties zoals ESHRE kunnen een centrale rol spelen in implementatiestrategieën, waarin naast de ontwikkeling van richtlijnen ook een implementatiesectie kan worden opgenomen voor elke richtlijn. De effectiviteit van geselecteerde implementatiestrategieën kan vervolgens in de toekomst worden bestudeerd en aangepast.

In de loop der jaren is de definitie van herhaalde miskramen en de bijbehorende risicofactoren uitgebreid besproken. Er is geen pathofysiologisch bewijs dat onderscheid maakt tussen vrouwen met twee zwangerschapsverliezen en vrouwen met drie of meer miskramen, maar er is enig bewijs dat de kans op bepaalde geassocieerde factoren, zoals het antifosfolipidensyndroom en het drager zijn van structurele chromosomale afwijkingen, niet verschillend is. **Hoofdstuk 6** biedt inzicht in deze oorzaken en geassocieerde factoren. We hebben vrouwen met twee miskramen en vrouwen met drie of meer miskramen vergeleken, en in ons cohort hebben we vastgesteld dat de risicofactoren voor herhaalde miskramen met dezelfde frequentie voorkomen in beide groepen. Deze bevinding biedt een basis om twee miskramen als definitie voor herhaalde miskramen te hanteren. Aangezien de frequentie van risicofactoren

vergelijkbaar is tussen de twee groepen, kan worden overwogen om onderzoek naar herhaalde miskramen te starten vanaf twee miskramen, vooral omdat bekend is dat stellen met herhaalde miskramen psychologische last ervaren vanwege het niet kunnen voortplanten. Op deze manier kan tijdig ondersteunende zorg worden geboden.

Een van de meest gestelde vragen van stellen met herhaalde miskramen heeft betrekking op hun toekomst: "Zullen we een zwangerschap succesvol kunnen dragen die resulteert in de geboorte van een levend kind? En zo ja, hoe groot is onze kans hierop?" Deze vragen kunnen worden beantwoord met behulp van predictiemodellen. Voordat deze modellen in de praktijk kunnen worden gebruikt, is het van belang dat ze nauwkeurig worden ontwikkeld en zowel intern als extern worden gevalideerd. Er bestaan verschillende voorspellende modellen voor herhaalde miskramen die in de klinische praktijk worden gebruikt, maar ze zijn nog nooit kritisch beoordeeld. Daarom is in **hoofdstuk 7** een systematische review uitgevoerd om bestaande predictiemodellen voor herhaalde miskramen te vinden en kritisch te beoordelen. Naarmate de kennis over de ontwikkeling en validatie van predictiemodellen in de loop der jaren is toegenomen, zijn er tools gepubliceerd die als leidraad dienen voor het rapporteren van alle benodigde items en voor het kritisch beoordelen van het risico op vertekening (bias) in studies die predictiemodellen proberen te ontwikkelen. We hebben aangetoond dat de zeven geïncludeerde predictiemodellen de aanbevolen ontwikkelingsstappen niet hebben gevolgd, waaronder interne validatie. Geen van de studies heeft een berekening van de steekproefgrootte uitgevoerd. Een retrospectieve berekening van de steekproefgrootte toonde aan dat drie studies te klein waren voor het aantal opgenomen voorspellende parameters. Het is met name opmerkelijk dat de predictiemodellen die door de ESHRE worden aanbevolen voor klinisch gebruik, namelijk die van Brigham et al. en Lund et al., beide een hoog risico op vertekening hadden, onder andere vanwege onvoldoende steekproefgrootte en categorisatie van voorspellers. Andere beperkingen zijn het niet rapporteren van ontbrekende gegevens, wat kan leiden tot selectiebias, en het opnemen van patiënten die zijn behandeld met verschillende therapeutische opties, wat van invloed kan zijn op het voorspellende effect van geselecteerde voorspellers.

Doorgaans worden de voorspellende prestaties van predictiemodellen beschreven aan de hand van discriminatie en kalibratie. Slechts twee studies rapporteerden de c-statistiek, een maat voor het onderscheidend vermogen van het voorspellende model. Geen van de geïncludeerde studies in dit hoofdstuk voerde kalibratie van hun model uit, wat mogelijk belangrijker is dan het beoordelen van het onderscheidend vermogen, omdat kalibratie de mate van overeenstemming van de gegevens met het ontwikkelde model beoordeelt.

Voordat klinische voorspellingsmodellen in de praktijk worden gebruikt, moet zowel interne als externe validatie worden uitgevoerd. Interne validatie is belangrijk omdat van het model wordt verwacht dat het goed presteert op zijn eigen gegevens, aangezien het model is ontworpen om bij de ontwikkelingsgegevens te passen. Zonder interne validatie van een voorspellend model kan er overfitting optreden, wat betekent dat het model minder nauwkeurig zal zijn bij testen in een nieuwe populatie met herhaalde miskramen. Externe validatie lijkt op interne validatie, maar wordt uitgevoerd in een niet-gerelateerd cohort. Het is nodig om te bevestigen dat het ontwikkelde predictiemodel de uitkomst voldoende goed kan voorspellen bij niet-gerelateerde patiënten in vergelijking met het cohort waarop het model is ontwikkeld. Als externe validatie wordt uitgevoerd en het voorspellingsmodel blijkt goed te presteren, kan overwogen worden om het model te implementeren.

Op basis van de conclusies en adviezen die we hebben geformuleerd in hoofdstuk 7, is er een externe validatiestudie uitgevoerd op de meest gebruikte en aanbevolen voorspellingsmodellen voor herhaalde miskramen. In **hoofdstuk 8** hebben we geprobeerd de voorspellingsmodellen van Brigham et al. te vergelijken met een cohort van 739 vrouwen met vergelijkbare kenmerken als het cohort van Brigham (gemiddelde leeftijd van 33,1 jaar en een mediane van drie zwangerschapsverliezen bij intake). We hebben aangetoond dat het voorspelde gemiddelde zwangerschapssuccespercentage in het Brigham-model 9,8 procentpunten hoger lag dan het waargenomen zwangerschapssuccespercentage in de dataset (73,9% vs. 64,0%). De prestaties werden gemeten aan de hand van kalibratie en discriminatie, waarbij de kalibratie liet zien dat het model overschatting vertoonde en te extreme voorspellingen deed. Tevens zijn de geschatte risico's te extreem.

Het onderscheidend vermogen van het model bedroeg 0,55. Deze waarde kan getallen aannemen tussen 0,5 en 1,0, waarbij 0,5 duidt op voorspellingen op basis van puur toeval en 1,0 duidt op perfecte voorspelling. Kortom, het model overschat, maakt te extreme voorspellingen en heeft een slecht onderscheidend vermogen.

De resultaten van de externe validatie van het voorspellingsmodel ontwikkeld door Brigham et al. kunnen worden beschouwd als een gevolg van het hoge risico op bias dat in **hoofdstuk 7** is ontdekt. De nauwkeurigheid van voorspellingsmodellen is vaak lager in een niet-verwant cohort, en dit geldt ook voor het model van Brigham. Een klein cohort van herhaalde miskramen patiënten en het ontbreken van interne en externe validatie hebben waarschijnlijk geleid tot de slechte prestaties van dit model. Zelfs na modelaanpassing door middel van herkalibratie, waarbij alle gebruikte coëfficiënten opnieuw worden geschat, was geen verbetering mogelijk. De resultaten van deze studie suggereren dat het momenteel meest gebruikte voorspellingsmodel voor stellen met herhaalde miskramen niet in staat is om onderscheid te maken tussen patiënten met wel of zonder een succesvolle toekomstige zwangerschap.

De huidige modellen baseren hun prognose alleen op twee voorspellers: het aantal eerdere miskramen en de leeftijd van de moeder. Dit leidt tot de vraag of de voorspellende capaciteit van het model zal verbeteren wanneer rekening wordt gehouden met aanvullende potentiële voorspellers. Er zijn echter weinig predictor studies uitgevoerd in herhaalde miskramen, en het selecteren van predictoren voor opname in predictiemodellen gebeurt vaak op theoretische gronden. Dit is momenteel de beste aanpak in afwezigheid van dergelijke predictor studies.

De vraag hoe goed een herhaalde miskramen voorspellingsmodel moet zijn om onderscheid te maken tussen degenen met en zonder het gewenste resultaat kan op verschillende manieren worden beantwoord. Een benadering is gebaseerd op de mogelijke waarden die een c-statistiek kan hebben, waarbij geldt dat een hogere c-statistiek beter is. Echter, predictor studies in de zwangerschap hebben vaak c-statistieken lager dan 0.65. Het is bekend dat zwangerschap een multifactorieel proces is dat nog niet volledig wordt begrepen in termen van wat leidt tot zwangerschapssucces en wat leidt tot het doormaken van een miskraam. Daarom is het

twijfelachtig of een predictiemodel ooit een hoger onderscheidend vermogen zal hebben dan wat momenteel in de literatuur wordt gevonden.

Hoofdstuk 9 presenteert een protocol voor de ontwikkeling en validatie van een nieuw predictiemodel voor herhaalde miskramen, met als doel de kans op een succesvolle zwangerschap in de toekomst bij stellen met herhaalde miskramen nauwkeurig te voorspellen. Het model richt zich zowel op stellen met als zonder onderliggende risicofactoren. Het primaire doel van het model is om de kans op een succesvolle geboorte binnen drie jaar na de eerste consultatie te voorspellen, wat klinisch relevanter is voor patiënten dan alleen de uitkomst van de eerste zwangerschap na intake. Daarnaast streeft het model ernaar om dynamisch de kans op een succesvolle geboorte te voorspellen op basis van elke uitkomst van de zwangerschap na de intake.

Bij de ontwikkeling en rapportage van het model is gebruik gemaakt van de TRIPOD-verklaring, die richtlijnen biedt voor het rapporteren van voorspellende modellen, en de PROBAST-tool, die helpt bij het minimaliseren van het risico op bias in het onderzoek. Door belanghebbenden, zoals klinici en patiëntenorganisaties voor stellen met vruchtbaarheidsproblemen, te betrekken, wordt gestreefd naar een solide basis voor het gebruik van dit model in de toekomst.

Het uiteindelijke doel van dit model is om op maat gemaakte en individuele schattingen te bieden van de kans op een succesvolle zwangerschap. Het model zal voortdurend worden bijgewerkt op basis van het nieuwste wetenschappelijke bewijs, zodat het de best mogelijke geïndividualiseerde counseling kan blijven bieden. Het is belangrijk om patiënten en klinici actief te betrekken bij de ontwikkeling en implementatie van dit predictiemodel om ervoor te zorgen dat het waardevol en bruikbaar is in de klinische praktijk.

CONCLUSIES EN BLIK OP DE TOEKOMST

Dit proefschrift heeft waardevol inzicht geboden in de dagelijkse praktijk van herhaalde miskramen zorg en counseling. Het onderzoek toont aan dat er ruimte is voor verbetering in de zorg en begeleiding van stellen met herhaalde miskramen. Het feit dat veel van deze stellen geen identificeerbare risicofactoren lijken te hebben, kan ontmoedigend zijn. Het is belangrijk om de variatie in praktijken en de redenen daarachter te erkennen en te streven naar effectieve, op bewijs gebaseerde zorg die het principe van "niet schaden" handhaaft.

De focus van toekomstig onderzoek moet liggen op het verbeteren van de kwaliteit van zorg en het beantwoorden van de meest dringende vragen in de herhaalde miskramen zorg. Begeleiding bij toekomstige zwangerschappen speelt een cruciale rol, zelfs in afwezigheid van effectieve behandelingsopties die de zwangerschapskansen verbeteren. Het is essentieel om goed ontwikkelde en gevalideerde predictiemodellen te gebruiken om paren te helpen weloverwogen beslissingen te nemen over verdere zwangerschapspogingen.

Samenwerking en verbinding tussen samenlevingen, onderzoekers en klinici over de hele wereld is van groot belang om de zorg voor herhaalde miskramen te optimaliseren. Met geavanceerde technologieën en wereldwijde connectiviteit kunnen we een nieuwe dimensie geven aan het onderzoek naar herhaalde miskramen en uiteindelijk de mysteries ervan ontrafelen. Door hoogwaardig onderzoek met een grote impact te bevorderen, kunnen we streven naar betere zorg en counseling voor stellen met herhaalde miskramen.