WHITEPAPER

Built for approval, not access

Why women's health R&D fails twice and how to fix it

By Katherine Broecker, Alejandra Manrique, Sneha Ramachandran and Jessica Williams

August 2026

Impact where it matters.®

Drug trials for women's therapies aren't designed with access in mind

Executive summary

Women spend 25% more of their lives in poor health than men. Despite representing half the global population and accounting for a disproportionate share of the chronic disease burden, women remain systematically underserved by the pharmaceutical and medical device industries intended to serve them.

The evidence for this pattern is consistent. Across five recent women's health drug approvals, regulatory success did not predict commercial access. In every case where reimbursement was constrained, the specific evidence gaps that payers cited were knowable and addressable before phase 3 enrollment began, including missing comparative effectiveness data, absent subgroup analyses, no safety data in breastfeeding women and delivery formats incompatible with real-world use. These evidence gaps were likely not addressed because the trials were not designed with access in mind. Some issues with these trials included:

- **Missing data for the population the drug is meant to treat.** Zulresso (brexanolone), approved for postpartum depression, required a 60-hour inpatient intravenous infusion that would separate mothers from their newborns. It also launched with no breastfeeding safety data.
- **No head-to-head comparison test against the standard of care.** Veozah (fezolinetant), a nonhormonal option for menopausal hot flashes, was never tested against hormone therapy, the established first-line treatment. Orilissa (elagolix) had the same gap. Its trials were placebo-controlled with no comparison to the hormonal treatments already used for endometriosis pain. Without head-to-head data, payers have little basis to justify covering either drug over the cheaper, established alternative.
- **A confirmatory trial that didn't reflect the population the drug was approved for.** Makena reached the market in 2011 through the U.S. Food and Drug Administration's (FDA) accelerated approval pathway based on a smaller trial that included 59% Black women, the population most affected by preterm birth. The confirmatory trial required to verify the results ran years later, with limited representation from Black women, failing to reproduce the original results. The drug was pulled from the market as a result, potentially cutting access for the very population it could have helped most.

The commercial upside of addressing trial design that creates evidence to support access is clear. When sex-specific trial design is incorporated at the start and demographic inclusion is a priority, the returns can be transformative. Novartis' Entresto (sacubitril-valsartan) demonstrates this directly in the cardiovascular space: A prospective sex-specific subgroup analysis supported an expanded FDA label, making 1.8 million more patients eligible for treatment.

We argue for a unified strategy that incorporates access challenges into R&D design from day one. Drawing on eight case studies, a structured analysis of barriers women face in clinical research, and a practical framework for R&D leaders and payers, we show that the regulatory and access challenges in women's health are not separate problems. They are compounding, structurally linked challenges with a common root cause and a common solution.

The case for a different approach to women's health R&D

Women's health has long been treated as a specialty—a niche within medicine defined primarily by reproductive function. The consequences of that framing have been profound. It has meant that conditions affecting women uniquely, differently or disproportionately have received less research attention, less funding and less clinical sophistication than conditions that affect both sexes or that have historically been associated with men.

The scale of this oversight is well documented. According to an analysis by the World Economic Forum (WEF) and McKinsey Health Institute, the global women's health gap represents 75 million disability-adjusted life years (DALY) annually, the equivalent of seven lost healthy days for every woman, every year. Women spend 25% more of their lives in poor health or with some degree of disability compared to men. Closing this gap could unlock at least $1 trillion in annual global GDP by 2040.

Despite this, only 7% of biopharma innovation is invested in women-specific conditions, and less than 1% of that investment goes to conditions beyond women's cancers. The National Institutes of Health (NIH) allocates just 11% of its budget to women's health research. Between 2019 and 2023, 11 startups addressing erectile dysfunction secured $1.24 billion in venture funding, while eight startups addressing endometriosis, a condition affecting an estimated 190 million women globally, received $44 million.

This underinvestment creates a compounding cycle. Limited funding produces limited science. Limited science produces trials designed to the minimum standard required for regulatory approval, with no systematic consideration of whether the evidence generated will satisfy payers, survive confirmatory scrutiny or reflect the populations who use the drugs. Regulatory approval is achieved. Access underperforms. ROI disappoints. Investment declines further.

We're not arguing that more should be spent on women's health. We're saying that what the industry does spend on must be designed differently, from a trial's conception through its enrollment strategy, evidence generation and the drug's commercial launch. The impact of better-designed R&D includes stronger evidence with commercial and clinical returns.

Part 1: The scale of the problem—a gap in women's health that costs everyone

The health burden for women

A common defense of the status quo in women's health research is that women live longer than men. While true, longevity alone does not reflect women's overall health experience.

A longer lifespan doesn't mean better health. It can mean more years spent sick. The health burden measure that captures this most directly is the DALY, a metric that combines years of life lost to premature death with years of life lived with disability or poor health. When the women's health gap is measured in DALYs, the picture is unambiguous. Women experience 75 million more DALYs per year compared to men, a disparity that we cannot explain by biological differences alone and that reflects structural failures in research, care delivery and investment.

The economic consequences are significant. Nearly half of this health burden falls on women of working age, the years when health has the most direct impact on economic participation,

productivity and earning potential. Addressing the gap could generate the equivalent impact of 137 million women accessing full-time employment positions by 2040 and could boost the global economy by at least $1 trillion annually. That figure is, if anything, an underestimate. It's based on data that systematically undercounts the prevalence of conditions like endometriosis and menopause and undervalues the disability burden of conditions that aren't fatal. Instead, these conditions chronically diminish women's capacity to work, care for their families and participate fully in life.

The approval-access disconnect in therapies for women

The dominant commercial model in pharmaceutical R&D treats regulatory approval as the primary milestone. The logic is intuitive: Without approval, there is no product. Without a product, there is no revenue. Everything is optimized to maximize the likelihood and speed of approval.

In most therapeutic areas, this model works reasonably well. In women's health, it does not. The evidence from recent approvals is consistent: Regulatory success in women's health does not predict success in commercial access. Drugs that clear FDA review can still face immediate, severe and sometimes permanent reimbursement constraints because of evidence gaps that existed in the trial data from the beginning.

This is the approval-access disconnect, and it has a specific structure. The evidence gaps that payers cite when restricting coverage—including missing comparative effectiveness data, no subgroup analyses in the highest-need populations, absent safety data in breastfeeding women and delivery formats incompatible with real-world clinical practice—do not emerge at launch. They are predictable requirements that payers communicate, sometimes explicitly and sometimes through their existing coverage frameworks, well before phase 3 enrollment begins. The gaps persist because trials are not designed with these requirements in mind.

There is also a second, less visible challenge. Some drugs achieve both regulatory approval and broad commercial coverage, only to collapse entirely when confirmatory trial evidence fails to support the original approval data. This is a predictable consequence of designing confirmatory trials for enrollment convenience rather than population validity. When the population in the confirmatory trial does not match the population in the original approval trial or the population using the drug, the evidence is fragile from the start.

Both challenges share a root cause: Trials are designed to satisfy regulators, not to build the durable, population-representative evidence base that sustained commercial access requires. Fixing both requires the same intervention: Integrating access into the design of development programs from the earliest stages, not as a late-stage commercial consideration but as a core scientific requirement.

Part 2: Missing from the beginning—the research gap in women's health R&D

A research system built around male biology

Modern biomedical research developed when the human body, for scientific purposes, meant the male body. Most animal models historically used in preclinical research have been male. The landmark clinical trials that established the evidence base for cardiovascular medicine, pharmacology and much of internal medicine enrolled primarily or exclusively men. Women's symptoms and clinical presentations that differed from the male norm were

not recognized as reflections of female biology; they were labeled “atypical,” a framing that located the deviation in the patient rather than in the research design.

The consequences of this framing extend beyond academic concern. Of 183 widely used medical interventions studied across 64 health conditions representing approximately 90% of the global disease burden in women, only 50% reported sex-disaggregated data. Among those that did, 64% of interventions were found to disadvantage women through lower efficacy, reduced access or both, compared to just 10% of interventions that disadvantaged men. Women in the U.S. have reported serious or fatal adverse events from approved medicines 36% more frequently than men since 2000. Since 1980, drugs have been withdrawn from the market for safety reasons at three-and-a-half times the rate in women compared to men.

These are not statistical artifacts. They are the predictable output of a research system that used the male body as its default template. Either the research system did not ask or it did not adequately study whether the same biology, the same treatments and the same dosing apply to women.

Clinical trial exclusion: The thalidomide legacy

The specific mechanism by which women were most systematically excluded from clinical research has a precise historical origin. In 1977, following the thalidomide tragedy of the 1950s and 1960s in which a sedative prescribed to pregnant women caused severe birth defects, the FDA issued guidance excluding women of childbearing potential from phase 1 and phase 2 clinical trials, except in cases of life-threatening conditions. The intent was protective and similar to early concerns in pediatric research. Regulatory mandates have transformed many of the pediatric concerns. Women, however, are not children, nor are they a rare population. The effect was to create a near-total data void on how the drugs developed over the following four decades behave in female bodies.

This guidance was reversed in 1993, but the void has not been filled. Women remain underrepresented in early stage trials. Women account for fewer than 30% of participants in industry-sponsored phase 1 trials, the phase at which dosing, pharmacokinetics and initial safety signals are established. The subpopulations most systematically excluded are women of color, post-menopausal women, and pregnant and lactating women. Just 5% of available medications have been adequately monitored, tested and labeled for use in pregnant and breastfeeding women, yet more than 80% of pregnant patients are routinely prescribed therapies that have never been studied during pregnancy or lactation.

The downstream consequences are significant and specific. When dosing is established in male-dominant phase 1 trials, women receive doses calibrated to male pharmacokinetics, metabolism and body composition. When safety signals are not disaggregated by sex, sex-specific adverse events are masked in aggregate data and do not appear in labeling. When pregnant and lactating women are excluded from trials, clinicians prescribing to this population do so without evidence. And payers covering these prescriptions do so without coverage frameworks that can distinguish appropriate from inappropriate use.

Missing data disaggregation in clinical trials

The exclusion problem is compounded by a reporting problem. Even in trials that enroll women in nominally adequate numbers, sex-disaggregated results are rarely reported. Across 52 migraine trials, only two trials published sex-disaggregated results, despite migraines being a condition that affects women at nearly three times the rate of men. Across 320 ischemic heart disease trials open to all sexes, only 26 trials published sex-disaggregated results, despite cardiovascular disease being the leading cause of death in women globally. Only 13% of Alzheimer's disease articles report sex-stratified findings, despite women representing approximately two-thirds of Alzheimer's patients.

The result is a systematic pattern in which aggregate trial data masks sex-specific effects, positive and negative, that would be visible if the data were routinely disaggregated. Payers who evaluate evidence packages built without disaggregated data cannot identify sex-specific subgroups who benefit most or are at greatest risk, which limits their ability to write coverage frameworks that support appropriate use. Clinicians who prescribe based on disaggregated label data are working with evidence that may not apply to the specific patient in front of them. And sponsors who skip analyzing sex-specific effects at the time of development cannot demonstrate the full value of their products to payers—as we will explain in the Entresto case—the expanded commercial opportunities that sex-specific analyses can reveal.

The pipeline misalignment for women's therapies

Investment and evidence gaps interact with a structural misalignment in how research priorities are set. Global life sciences R&D has historically prioritized conditions with high years-of-life-lost contributions or conditions that are life-threatening. This is a rational prioritization in a world where mortality is the primary measure of disease burden and where regulatory pathways and payer frameworks are calibrated to mortality-anchored endpoints.

The problem for women's health is that many of the conditions causing the greatest suffering among women are not fatal. Endometriosis causes severe and often debilitating chronic pain, infertility and years of diagnostic delay, but it does not appear in most mortality statistics. Menopause affects virtually every person who is female at birth and reaches midlife, causing symptoms that WEF research indicates reduce economic productivity by up to $120 billion annually in GDP terms. But the Global Burden of Disease dataset doesn't even give it a standalone category. Premenstrual syndrome affects tens of millions of women during their working years and has a larger estimated economic impact than ischemic heart disease in the WEF analysis, but it receives a fraction of the research investment.

When the prioritization framework is built around mortality and years-of-life-lost, and when the conditions causing the greatest burden in women are predominantly disability-causing rather than life-shortening, the pipeline will systematically underinvest in women's health relative to its true burden. Addressing this requires more than increasing the volume of investment. It requires changing the metrics by which conditions are assessed for research priority and changing the payer value frameworks that give those metrics their commercial consequences.

Part 3: The structural barriers to women's trial participation

Barriers that compromise clinical trial enrollment and data quality

The design of clinical trials has historically optimized for regulatory approval, not for the population that will ultimately use the approved product. In women's health, this misalignment creates a consistent set of structural barriers that reduce enrollment, compromise population representativeness and produce evidence that is weaker and more vulnerable to payer challenges than it needs to be.

These are the most consequential barriers to trial enrollment that have downstream effects on evidence quality and payer access:

- **Scheduling and caregiving burdens.** Clinical trials based on regular site visits assume participants can reliably travel to academic medical centers, take time off work and arrange child care. Women, who continue to bear a disproportionate share of caregiving responsibilities, are more likely to be unable to meet these requirements. The result is underenrollment from populations that are already underrepresented in trials—working-age women, women with young children and women in lower-income groups. The populations excluded are often the ones with the highest unmet need and the ones payers will most scrutinize for coverage purposes.

- **Geographic concentration.** Academic medical center-based trial sites concentrate geographically in urban, higher-income areas, and trials often recruit from the patient populations associated with those institutions. In women's health, where conditions like endometriosis are significantly underdiagnosed in lower-income and underserved populations, this geographic concentration creates systematic gaps in population representativeness that are visible in the data and flagged in payer assessments.

- **Hormonal variability exclusion.** Phase 1 trials, which establish the dosing, pharmacokinetics and early safety profile that inform all subsequent development, have historically enrolled predominantly male participants. This is, in part, because researchers have treated hormonal variation as a confounding variable rather than a biological feature to be understood. The practical consequence is that dose-finding is conducted in male biology and women receive doses calibrated to male metabolism. This creates both safety risks—sex-specific adverse events that are not detected until post-market—and payer exposure. This is because underpowered female adverse drug reaction data generates coverage uncertainty.

- **Pregnancy and lactation exclusions.** The near-total exclusion of pregnant and lactating women from clinical trials, a legacy of the 1977 FDA guidance, means that approximately 80% of pregnant patients are prescribed drugs with no pregnancy-specific safety or efficacy data. This creates a specific payer consequence: In therapeutic areas where the target population includes women of reproductive age, postpartum depression being the clearest example, the absence of data in breastfeeding women becomes a primary coverage restriction driver.

- **Patient-reported outcome instrument misalignment.** The outcome measures used in women's health trials are often validated in populations that do not match the trial's target population. Payers increasingly require health-related quality-of-life evidence to justify premium pricing, particularly in conditions they classify as quality-of-life indications.

When assessments of patient-reported outcomes are not prespecified, validated in the target population and aligned with payer value frameworks, the health-related quality-of-life evidence generated is often insufficient for coverage purposes.

- **Menstrual cycle data.** Phase-dependent variation in pharmacokinetics, symptom severity and treatment response is well documented for a range of conditions including migraine, epilepsy, asthma and depression. Yet the menstrual cycle phase is routinely uncollected in clinical trials, which can cause missed safety signals when adverse events are phase-dependent and leave post-marketing surveillance gaps after approval.

The women-as-an-afterthought trial design problem

The individual barriers described above are symptoms of a more fundamental problem: Clinical trial design has treated women's participation as an accommodation to the standard male-centered design, rather than as a primary design requirement.

When women are added to trials designed around male physiology, male schedules and male-typical disease presentations, the result is data of lower quality than would come from trials designed with women's participation as a primary constraint. The Renovia case, discussed in detail in our case evidence, demonstrates this directly. A trial originally planned as a 15-site, site-based study was converted to a fully virtual, decentralized design during the COVID-19 pandemic. The virtual design produced better data, higher retention and a more diverse enrolled population than the site-based design would have achieved. Women-centered design was not a constraint on evidence quality—it was the mechanism that produced it.

Medical device development provides the starkest physical illustration of what happens when male physiology is treated as the default. Some hip implants fail at nearly twice the rate in women as in men due to differences in bone structure and load-bearing mechanics that were not incorporated into the original device design. The same logic applies to drug development. Dosing recommendations derived from male-dominant pharmacokinetic studies carry analogous risks: incorrect dosing, underdetected adverse events and label restrictions that reduce access for the populations with the highest need.

The payer value framework problem

There is a third structural barrier that is distinct from both the research gap and the trial design problem, and that is specific to women's health—the systematic undervaluation of the conditions that disproportionately affect women within standard payer value frameworks.

Many of the conditions causing the greatest burden in women, including genitourinary syndrome of menopause, premenstrual syndrome and endometriosis, payers categorize as quality-of-life conditions rather than clinical disease states. This categorization is not neutral. It triggers a different reimbursement logic. For example, this logic leads to higher evidence thresholds for coverage, default step-therapy requirements, resistance to premium pricing and a cost-effectiveness framework anchored on mortality outcomes that systematically undervalue disability-causing conditions.

The practical consequence is that even well-designed trials generating robust efficacy data can fail to achieve access if the condition being treated is not valued appropriately in the payer framework. The problem here is architectural, not due to evidence quality. And it is one that the industry can only address through proactive payer engagement before trial design begins, not through post-approval negotiation.

Part 4: The case evidence: 8 studies in approval, access and design

Illustrations of the approval-access challenge

The eight case studies highlighted here are drawn from recent women's health product development. Together they represent the full spectrum of the approval-access challenge: drugs that were approved but stranded commercially, a drug whose evidence collapsed under confirmatory scrutiny and two products whose design choices produced meaningfully better clinical and commercial outcomes.

The cases are organized into two groups: six underperformers and two successes. The taxonomy is analytical, not evaluative. And the underperforming cases are presented to identify the specific design decisions that produced poor access outcomes, not to criticize the sponsors involved. In most cases, the evidence gaps that drove reimbursement failure were predictable from the structure of the trial, not from hindsight.

> In every underperforming case, the evidence gaps that drove disappointing outcomes were knowable before phase 3 enrollment began. In both success cases, deliberate choices at the design stage produced better evidence and better commercial results. The difference between the two groups is not post-market luck. It's upstream design intent.

6 underperforming cases

Case 1: Zulresso (brexanolone) and delivery format issues

In March 2019, brexanolone became the first drug approved by the FDA specifically for the treatment of postpartum depression in adults—marking a genuine clinical advance. It was a neuroactive steroid that modulated GABA-A receptors through a mechanism directly linked to the hormonal changes of the postpartum period. Phase 3 trial data showed a significant reduction in depressive symptoms within 60 hours. But the drug was commercially stranded almost immediately.

The mechanism was a 60-hour continuous intravenous infusion, administered in a certified healthcare facility, with continuous pulse oximetry monitoring and a healthcare provider available throughout. Patients had to be monitored every two hours during waking hours. The risk evaluation and mitigation strategy (REMS) program required facility certification, pharmacy certification and patient enrollment before each infusion. By December 2021, two-and-a-half years after FDA approval, only 499 patients had received treatment in the postmarketing surveillance program.

The delivery format was not a post-approval surprise. The clinical rationale for IV administration was understood from the molecule's earliest development. The REMS requirement followed directly from the sedation risk inherent to the mechanism of action. Both were visible, predictable and addressable at the design stage, through formulation work that would have prioritized an oral or subcutaneous delivery format, or through a trial design that specifically measured the feasibility of the delivery format in real-world postpartum settings.

The breastfeeding data gap compounded the access problem. Postpartum depression occurs, by definition, in women who have recently given birth, many of whom are

breastfeeding. Brexanolone passes into breast milk. The label contained no safety data in breastfeeding women, because breastfeeding women were excluded from the trials. For a condition that is specific to new motherhood, the absence of breastfeeding data was not a peripheral gap. It was a primary access barrier, because payers writing coverage criteria for a new postpartum drug will always ask whether the drug is safe for the infant.

Zulresso was withdrawn from the U.S. market in April 2025 for business reasons, a commercially quiet end to a clinically important story. The lesson is not that the science was wrong. It's that a mechanistically validated molecule was stranded by a delivery format and population exclusion that were foreseeable from the earliest stages of development.

Case 2: Zurzuvae (zuranolone) and incomplete design improvement

Zuranolone, approved by the FDA in August 2023 and marketed as Zurzuvae, represented a meaningful step forward from brexanolone. It was the first oral treatment specifically approved for postpartum depression, a 14-day course of once-daily capsules that could be taken at home. The delivery format issue had been solved. The access issue had not.

The FDA approved zuranolone for postpartum depression but issued a Complete Response Letter (CRL) for its major depressive disorder (MDD) indication, finding that the evidence did not demonstrate substantial effectiveness for MDD. The rejected MDD label was not an incidental omission; it was the commercial thesis. Postpartum depression affects approximately one in seven new mothers while MDD affects tens of millions more. The MDD CRL narrowed the addressable market significantly and left zuranolone with an indication whose patient population is defined by a time-limited postpartum window rather than a chronic condition requiring long-term treatment.

The breastfeeding data gap, inherited directly from brexanolone, persisted. Participants in the zuranolone clinical trials were asked to refrain from breastfeeding during the trial. No data was generated on secretion into breast milk or effects on breastfed infants. For postpartum women who are breastfeeding—a significant portion of the population the drug is designed to treat—the absence of data creates both clinical uncertainty and payer coverage restrictions.

Sage Therapeutics had initially aimed to price zuranolone below $10,000. When the MDD CRL was issued, the commercial basis for that pricing was weakened and the final wholesale acquisition cost was set at $15,900 for a 14-day treatment course, established after the evidence package was locked rather than designed to reflect a prespecified value story. Early coverage analysis found that fewer than 1% of the nation's approximately 1,000 private insurers had published coverage guidelines in the weeks following launch. Step-therapy requirements became common among those that did.

By December 2024, approximately 95% of commercial and Medicaid patients had nominal coverage, but coverage was subject to prior authorization, and step-therapy requirements remained the norm. The incremental improvement over brexanolone was real. The access strategy, designed after the evidence was locked, remained reactive.

Case 3: Veozah (fezolinetant) and a missing comparator

Fezolinetant, approved by the FDA in May 2023 and marketed as Veozah, addressed one of the most significant unmet needs in women's health: moderate to severe vasomotor symptoms, hot flashes and night sweats, in menopausal women who cannot or choose not to use hormone therapy. It was the first nonhormonal neurokinin-3 receptor antagonist

approved for this indication, offering a mechanistically novel alternative for a population with limited options.

As of July 2024, approximately 64% of commercial covered lives had coverage for Veozah, with prior authorization and step-therapy requirements.

The primary payer friction point was a gap entirely predictable from the structure of the clinical program. Hormone therapy remains the standard of care for vasomotor symptoms of menopause. It is effective, widely prescribed and inexpensive in generic form. Payers evaluating a new, nonhormonal treatment at a premium price want to know one thing above all others: How does it compare to the existing standard of care? The SKYLIGHT trials, the phase 3 program supporting Veozah's approval, were placebo-controlled. There was no head-to-head comparison with hormone therapy.

The subgroup gap reinforced the coverage friction. The commercial and clinical rationale for fezolinetant is strongest in women who cannot use hormone therapy, women with hormone-sensitive cancers and women with a history of blood clots—women for whom hormonal treatment carries unacceptable risks. This is the population with the highest unmet need and the one for whom fezolinetant's nonhormonal mechanism is most directly relevant. Yet this population, the menopausal-hormone-therapy (MHT) ineligible subgroup, was not studied as a primary subgroup with its own powered analysis. The evidence was weakest precisely where the value proposition was strongest.

Case 4: Orilissa (elagolix) and chronic access friction

Elagolix, approved by the FDA in July 2018 and marketed as Orilissa, was the first new treatment for endometriosis-associated pain in more than a decade. It addressed a condition affecting an estimated 190 million women globally, with an average diagnostic delay of approximately 10 years and no available cure. The clinical need was real and large.

Seven years after approval, elagolix remains chronically constrained commercially. Prior authorization is required for 70% to 80% of commercial and Affordable Care Act (ACA) marketplace insurance plans, and approximately 50% of Medicaid plans. Step therapy, requiring patients to fail other treatments before elagolix is covered, is required for approximately 50% of commercial, ACA and Medicaid plans. The manufacturer's co-pay assistance program explicitly excludes Medicaid, Medicare and other government-funded insurance programs, a particularly consequential restriction given that Medicaid covers a disproportionate share of reproductive-age women, many of whom are in lower-income groups whose endometriosis is most likely to go undiagnosed or undertreated.

The primary evidence gap driving this persistent friction is the absence of an active comparator trial. Payers evaluating a new hormonal treatment for a chronic condition with established therapeutic alternatives, including GnRH agonists, progestins and oral contraceptives, expect comparative effectiveness data. The pivotal elagolix trials were placebo-controlled, leaving no direct evidence of comparative efficacy against existing standards of care. The label also includes treatment duration restrictions of six months at the higher dose and 24 months at the lower dose, introduced in response to concerns about bone mineral density. These restrictions add further complexity to coverage decisions because chronic pain conditions like endometriosis typically require long-term management.

The Medicaid exclusion from co-pay assistance deserves particular attention. The population most affected by endometriosis-associated pain includes a significant proportion of

lower-income women who are more likely to experience diagnostic delays. Women in this population are also more likely to have severe, undertreated disease at diagnosis and are more likely to be covered by Medicaid. Designing a commercial access strategy that excludes this population is a design choice with foreseeable equity consequences.

Case 5: Osphena (ospemifene) and Intrarosa (prasterone) and payer value framework challenges

Ospemifene and prasterone are both FDA-approved treatments for genitourinary syndrome of menopause (GSM), a condition characterized by vulvovaginal atrophy, vaginal dryness, dyspareunia and urinary symptoms that affect most postmenopausal women and worsen without treatment. Both have clinical efficacy data. Neither faces a primary evidence quality problem. Both face persistent commercial access constraints.

The difference between this case and the preceding four is important. Its challenges do not result from trial design, but rather from market architecture.

Payers have systematically categorized GSM as a quality-of-life condition rather than a clinical disease state. This categorization triggers a different reimbursement logic than disease-state coverage: higher evidence thresholds, default preference for over-the-counter alternatives and resistance to premium pricing. It also relies on a cost-effectiveness framework anchored to mortality outcomes, which systematically undervalues conditions that aren't fatal but significantly impair quality of life, sexual health and urinary function.

Standard cost-effectiveness models are not well suited to valuing this burden. DALY calculations for GSM are relatively low compared to life-threatening conditions, not because the suffering is mild, but because the condition does not shorten life. When payer value frameworks use DALY-based cost-effectiveness as the primary criterion, conditions like GSM will consistently appear less valuable than they are in practice. Ospemifene and prasterone are trapped in a payer architecture that was not built with their condition category in mind.

Case 6: Makena (17-alpha hydroxyprogesterone caproate) and evidence collapse

The Makena story is the most instructive case in this paper because it's the inverse of the pattern in the preceding five. Rather than a drug that gained regulatory approval but failed to achieve commercial access, Makena achieved approval, broad reimbursement coverage, widespread clinical adoption and then collapsed entirely when the confirmatory evidence it should have generated from the beginning was not there.

Makena, the brand name for 17-alpha hydroxyprogesterone caproate, or 17-OHPC, received accelerated FDA approval in 2011 for the prevention of recurrent preterm birth in women with a history of spontaneous preterm delivery. The accelerated approval was based on a 2003 NIH-sponsored randomized controlled trial showing a roughly one-third reduction in preterm birth in treated women. The trial enrolled a population that was approximately 59% Black, which was appropriate given that Black women in the U.S. experience preterm birth at nearly twice the rate of white women.

The accelerated approval came with a requirement: Makena's sponsor must conduct a larger confirmatory trial demonstrating clinical benefit, not only in preterm birth rates but also in neonatal health outcomes. This confirmatory trial, the PROLONG study, enrolled 89% white participants, with 61% of participants recruited from Russia and Ukraine, and only 23% from the U.S. The population that bore the greatest burden of preterm birth in the U.S. was barely represented in the trial designed to confirm the drug's effectiveness.

The population mismatch between the original approval trial and the confirmatory study is not a scientific accident. It's the foreseeable consequence of designing a confirmatory trial for enrollment convenience and easier recruiting, rather than population validity. The U.S. population that used Makena and that bore the highest burden of preterm birth was not the population studied in the confirmatory trial. When the evidence did not hold, those women were left with nothing.

2 successful cases

Case 1: Renovia Leva as an example of women-centered decentralized trial design

Renovia is a digital therapeutics company that developed the Leva Pelvic Health System, an FDA-cleared prescription device combining a vaginal motion sensor with a smartphone app to guide pelvic floor muscle training for the treatment of urinary incontinence. In 2020, Renovia initiated a randomized controlled trial to generate definitive evidence of Leva's clinical efficacy that would support both FDA clearance and payer coverage engagement.

The trial was originally designed as a 15-site, site-based study. The COVID-19 pandemic disrupted that plan. Rather than pause the program, Renovia converted the trial to a fully virtual, decentralized design, the first fully virtual clinical trial in the urogynecology space. All 363 enrolled participants participated entirely from home, with no site visits required. Real-time data was captured through the Leva app, eliminating recall bias. A virtual site team monitored compliance, engaged with participants and escalated adverse events in real time.

The results were better than a site-based design would likely have produced. The trial enrolled a population that was 54.9% postmenopausal and 84.6% parous, exactly the demographics most affected by urinary incontinence, at participation rates higher than typical site-based trials achieve. Retention through six and 12 months was 95.7%. The published findings, appearing in "Obstetrics & Gynecology" in 2022, demonstrated significantly greater improvement in urinary incontinence symptoms in the Leva group compared to the standard home program at both six and 12 months.

The trial was recognized with the 2021 Medical Device Network Excellence Award for Research and Development. The Leva device later received FDA clearance for an expanded indication in fecal incontinence in July 2022, supported by additional evidence generated using the same decentralized clinical trial (DCT) methodology.

The Renovia case demonstrates something that the underperforming cases collectively obscure: Women-centered trial design is the mechanism that produces quality evidence. A design that reduces the participation burden for the target population enrolls a more representative sample, retains participants longer and generates data that more accurately reflects real-world use. The evidence is simultaneously more robust scientifically and more relevant commercially because it reflects the population payers will encounter in practice.

Case 2: Entresto (sacubitril-valsartan) demonstrating the commercial upside of sex-specific design

The Entresto case begins in 2015, when the FDA first approved sacubitril-valsartan for heart failure with reduced ejection fraction (HFrEF). This is a condition that had historically been studied predominantly in male patients and for which approved treatments had been validated in clinical trial populations that skewed heavily male.

The PARADIGM-HF trial, on which the original approval was based, enrolled approximately 22% women. Women are underrepresented in the HFrEF population, which is itself more

common in men, so this was not far from epidemiologically appropriate. The drug was approved and became one of the most commercially successful cardiovascular agents of its era.

What happened next is why we consider this case a success for women's health. Novartis conducted the PARAGON-HF trial to evaluate sacubitril-valsartan in heart failure with preserved ejection fraction (HFpEF), a condition that affects approximately 50% of heart failure patients and in which women are disproportionately represented, being twice as likely as men to develop HFpEF. The trial enrolled more than 4,800 patients and narrowly missed its primary endpoint overall, with a rate ratio of 0.87 that fell just short of the prespecified significance threshold.

In most development programs, a trial that misses its primary endpoint is a failure. In this case, prespecified subgroup analyses, including a sex-based subgroup that had been explicitly incorporated into the trial design, identified a significant treatment benefit in women, supported by a subsequent meta-analysis across both PARAGON-HF and PARADIGM-HF. Women with HFpEF showed consistent benefit from sacubitril-valsartan; the overall trial result was diluted by the male subgroup, in which the benefit was less pronounced.

The FDA granted an expanded indication in February 2021 based on this evidence, making Entresto the first drug in the U.S. indicated for chronic heart failure not specifically characterized by ejection fraction. The expanded label made the drug available to patients with HFpEF whose ejection fraction was "below normal," a category that, based on analysis of approximately 4.7 million U.S. adults living with heart failure, represented up to 1.8 million newly eligible patients with the potential to prevent or postpone as many as 180,000 worsening heart failure events.

The commercial implication is direct. The sex-specific subgroup analysis, built into the trial design prospectively and powered to detect the effect it found, produced an expanded commercial indication worth billions in additional revenue and clinical benefit for millions of patients, in a therapeutic area that had initially been studied without adequate attention to sex-based differences. The original HFrEF approval left a significant population of heart failure patients, disproportionately women, without an approved treatment. The subsequent sex-specific analysis found that treatment.

The pattern across all 8 cases

The eight cases described above represent a range of conditions, therapeutic mechanisms, delivery formats and commercial contexts. Yet they all share a consistent underlying pattern

In all six underperforming cases, the evidence gaps that drove poor access outcomes were predictable before phase 3 enrollment began. In both success cases, deliberate choices at the design stage produced better outcomes.

The difference between the underperformers and the successes is not primarily about resources, regulatory relationships or commercial sophistication. It's about design intent. The outcomes are different when:

- Access is built into R&D design from the earliest stages
- The question "what evidence will payers require?" is asked before phase 3 protocols are written

- "How do we enroll the population that actually uses this drug?" is treated as a scientific constraint rather than an enrollment logistics problem

Product	Type	Approval	Access	Key evidence gap	Access-oriented design
Zulresso (brexanolone)	Delivery format issue	FDA 2019	Very low (withdrawn 2025)	IV delivery, REMS no breastfeeding data	No
Zurzuvae (zuranolone)	Incomplete improvement	FDA 2023	Partial (~95%, step therapy)	MDD CRL, breastfeeding gap, reactive pricing	Partial
Veozah (fezolinetant)	Missing comparator	FDA 2023	~64% commercial, step therapy	No head-to-head vs. menopausal hormone therapy (MHT), no MHT-ineligible subgroup	No
Orilissa (elagolix)	Chronic access friction	FDA 2018	Chronic PA + step therapy, Medicaid gap	No comparator trial, duration limits, Medicaid exclusion	No
Intrarosa (prasterone) and Osphena (ospemifene)	Value framework gap	FDA approved	Systematic payer discounting	Condition categorized as quality-of-life, not disease state	No
Makena (17-OHPC)	Evidence collapse	FDA 2011	Broad → complete collapse 2023	Confirmatory trial population mismatch	No
Renovia Leva (DCT)	Design success	De Novo	Trial quality validated	N/A, evidence generation success story	Yes
Entresto (sacubitril-valsartan) (women subgroup)	Commercial success	FDA expanded 2021	1.8M new eligible patients	N/A, sex-specific design drove expanded label	Yes

FIGURE 1: Comparative examples of evidence-generation, access and commercialization outcomes for women's health products

Part 5 : The policy environment is progressing for women's health. Trial sponsors should take the lead.

The regulatory tailwind

The regulatory environment in the U.S. and internationally is shifting in ways that increasingly align with the access-oriented R&D model described in this paper. These changes are not yet sufficient to transform the field. Voluntary guidelines have consistently underperformed relative to requirements, but they represent a meaningful shift in the baseline expectation for trial design.

While FDA guidance has been shifting with the current administration, the FDA's 2024 final guidance on DCTs endorses DCT elements as standard practice rather than exceptional deviation, providing regulatory cover for sponsors who wish to use virtual or hybrid trial designs. The same year, the FDA issued its Diversity Action Plan guidance, requiring

sponsors to submit diversity action plans for phase 3 trials with specific inclusion targets, including for pregnant and lactating women. The International Council for Harmonization published ICH E21 in 2023, establishing an international framework for the inclusion of pregnant and breastfeeding individuals in clinical trials. The NIH's Sex as a Biological Variable policy, in effect since 2016, requires sex to be considered as a biological variable in all federally funded research, with direct implications for preclinical study design and clinical trial analysis.

Together, these policy developments signal that the research community is moving toward the practices this paper advocates. Sponsors who implement these practices now, not because they are required but because they produce better evidence, will have built institutional capability and regulatory credibility that positions them ahead of incoming mandates.

The incentive landscape

Policy is also beginning to address the funding gap that underlies much of the evidence quality problem. The WEF 2025 policy paper on women's health research, developed with input from more than 45 organizations including industry, regulators and patient groups, proposes a package of incentives modeled on the Orphan Drug Act, encompassing:

- Female disease designation
- Priority review vouchers
- Tax credits for clinical trial expenditures
- Public-private investment matching mechanisms

The pediatric precedent is directly instructive here. The Pediatric Research Equity Act, which mandated studies in children rather than incentivizing them voluntarily, produced 475 drug approvals, compared to 199 under the voluntary Best Pharmaceuticals for Children Act. The same logic applied to women's health suggests that mandatory inclusion and disaggregation requirements, combined with financial incentives for the highest-need conditions, would produce a step-change in evidence quality that voluntary approaches alone have not achieved.

Part 6 : A unified framework for designing women's health R&D for access

Our hypothesis

The eight cases in this paper support a single, testable hypothesis: Earlier integration of reimbursement into R&D decisions, combined with improved women's trial experience, can shift investment decisions and improve evidence quality in women's health.

The cases allow a stronger version of this hypothesis. Access-oriented design does not merely reduce reimbursement friction after the fact. It generates expanded commercial indications, as Entresto demonstrates. It produces more durable evidence that survives confirmatory scrutiny, as Makena's challenges illustrate in reverse. And it creates a feedback loop in which better evidence supports broader access, which supports stronger commercial returns, which supports reinvestment in the conditions that have been historically underserved.

The unified strategy has two interdependent pillars: reimbursement-centered design and women-centered trial design. Neither is sufficient alone. A trial that centers women's

participation but generates evidence that fails payer value requirements will not be successful. A trial that engages payers early but continues to exclude the highest-need populations will generate evidence that is commercially fragile. Both pillars must be present, and both must be incorporated from the earliest stages of development.

Recommendations for R&D leaders to design for access to women's therapies in mind

1. Integrate reimbursement before phase 3. The most important single change R&D leaders can make is to treat payer evidence requirements as an input to phase 3 design, not an output of the commercial launch process. This means conducting payer advisory engagements before phase 3 protocols are finalized, specifically to identify whether the condition is categorized as a disease state or quality-of-life indication, what comparative effectiveness data is required, which subgroups need powered analyses and what delivery format or administration requirements will create coverage barriers.

2. Treat population validity as a scientific requirement. The Makena case demonstrates the cost of designing for enrollment convenience at the expense of population validity. Every phase 3 trial in women's health should be able to answer one question before enrollment begins: Does the population we are proposing to study match the population that will use this drug? If the answer is no, the trial is building on a fragile evidentiary foundation, regardless of how large it is.

3. Prespecify sex as a primary subgroup variable. In any therapeutic area where women carry a meaningful share of the disease burden, sex should be specified as a primary subgroup variable in phase 2 and phase 3 trial designs, put in place to detect clinically meaningful sex-specific effects. The Entresto case shows a sex-specific signal can generate an expanded commercial indication worth billions. That signal is only actionable if the subgroup analysis was prespecified, not found retrospectively.

4. Design trials for the women who will use the product. This means:

- Applying decentralized trial elements to reduce scheduling and geographic barriers
- Presuming women of childbearing age eligible for participation rather than routinely excluding them
- Harmonizing contraception requirements based on actual reproductive risk rather than broad precautionary standards
- Collecting menstrual cycle phase data as standard
- Using patient-reported outcome instruments validated in the target population and aligned with payer value requirements

5. Build breastfeeding and pregnancy safety data into the development plan. For any indication where the target population includes women of reproductive age, especially postpartum conditions, the absence of breastfeeding safety data is a primary, predictable and avoidable access barrier. Planning for this data collection, modeled on the maternal investigation plan framework proposed in the WEF and Kearney policy paper, should be a standard element of phase 2 development plans.

Recommendations for payers and health technology assessment (HTA) bodies evaluating therapies for women

1. Publish evidence requirement frameworks earlier. The most common driver of evidence gaps in women's health trials is not that sponsors chose to generate inadequate evidence, it's that they weren't aware of what payers required when the trial was designed. Proactive publication of evidence frameworks for high-burden women's health conditions, modeled on the National Institute of Health and Care Excellence's early scientific advice program, would enable sponsors to design for coverage from the outset rather than discovering coverage requirements at launch.

2. Develop sex-specific value frameworks. Standard cost-effectiveness frameworks anchored on mortality outcomes systematically undervalue the disability burden of conditions that predominantly affect women. Developing sex-specific value frameworks, or condition-specific frameworks for indications like GSM, endometriosis and menopause that appropriately weight disability, productivity loss and quality of life would align payer value assessments more accurately with the actual burden of these conditions.

3. Distinguish evidence design failures from genuine scientific uncertainty. When reviewing coverage for newly approved women's health products, payers should distinguish between:

- Evidence gaps that reflect poor trial design
- Missing comparators
- Absent subgroup analyses
- Excluded populations and evidence gaps that reflect genuine scientific uncertainty that could not have been resolved at the time of development.

The evidence design failures are addressable through conditional coverage linked to post-market evidence generation; scientific uncertainty may require different approaches. Treating all evidence gaps as equivalent creates a coverage environment that penalizes avoidable design failures and unavoidable scientific uncertainty alike.

Part 7 : The investment case for women's health R&D

The argument for access-oriented R&D in women's health is not primarily a financial one. It's a clinical one: Women deserve research that reflects their biology, trials that enroll their populations and treatments that are accessible when they need them. That case stands independently of any commercial consideration.

But the financial case is also real. For audiences whose primary accountability is to investors, shareholders and health system budgets, it deserves to be made directly.

For every $1 invested in closing the women's health gap, approximately $3 is projected in economic returns globally, through improved workforce participation, reduced absenteeism and presenteeism, and decreased healthcare utilization. Endometriosis and menopause alone represent an estimated $130 billion in addressable GDP impact by 2040 once the systematic undercounting of their prevalence is corrected.

Access-oriented R&D is not more expensive than approval-first R&D. It requires different design choices, earlier payer engagement and a willingness to treat population validity as

a scientific requirement rather than an enrollment convenience. These are not primarily cost drivers. They are design priorities. And the evidence from both successes and underperformers suggest they are the design priorities that separate products with durable commercial access from products that achieve regulatory approval and then stall.

The window for women's health R&D is open

Women's health research is at an inflection point. The policy environment is shifting. Regulatory requirements for diversity and sex-specific analysis are strengthening. The commercial case for women-centered design is increasingly well evidenced. The underperformers are sufficiently numerous and sufficiently well documented that the pattern is impossible to dismiss as coincidence.

What remains is a choice about whether the R&D community will wait for requirements to arrive or begin designing for access now.

The case studies in this paper suggest that waiting has a cost. The evidence gaps driving reimbursement failure in women's health are predictable. They are addressable. They appear in trial after trial not because they are technically unavoidable but because the trials were not designed with access in mind. Each approval that fails to achieve commercial access is not just a commercial disappointment, it's a missed opportunity to generate the evidence that would support the next investment, the next trial and the next product in a therapeutic area that desperately needs them.

The successes suggest the alternative is achievable. Renovia designed a trial around women's lives and produced better evidence than a conventional trial would have generated. Novartis treated sex as a scientific question rather than a demographic variable and discovered an expanded commercial indication that the original development program had not been designed to find. In both cases, the design choice that produced better clinical evidence also produced better commercial outcomes.

The window is open. The regulatory environment supports women-centered design. The commercial upside is evident. The pattern of underperformance is documented. What the field needs now is a commitment, at the R&D design stage, to build trials that are not just built for approval but designed for access.

Women have waited long enough.

References

Research reports

- World Economic Forum in collaboration with the McKinsey Health Institute. (2024, January). Closing the Women's Health Gap: A $1 Trillion Opportunity to Improve Lives and Economies. PDF.
- World Economic Forum in collaboration with the Global Alliance for Women's Health, Kearney and the Gates Foundation. (2025, May). Prescription for Change: Policy Recommendations for Women's Health Research. PDF.
- ZS. (2025). Women's Health R&D: Built for Approval, Not Access. Internal analysis.

Case study sources: Zulresso (brexanolone)

- FDA. (2019). Zulresso (brexanolone) Prescribing Information.
- South Carolina Blues. (2025). Brexanolone (Zulresso) for Postpartum Depression. Medical Policy.
- Garafola et al. (2023). Safety of Brexanolone in Adults with Postpartum Depression: Postmarketing Surveillance Data. PMC.

Case study sources: Zulresso (brexanolone)

- Biogen/Sage Therapeutics. (2023 August 4). FDA Approves ZURZUVAE (zuranolone).
- Policy Center for Maternal Mental Health. (2024, December). One Year Later: How Insurers Are Covering Zurzuvae.
- MGH Center for Women's Mental Health. (2023). New Oral PPD Treatment Zuranolone Will Soon Hit the Market as Zurzuvae.
- Drug Discovery Trends. (2023 November 7). Is the new $15,900 postpartum depression pill worth it?
- KFF Health News and NPR Health. (2024 March 12). A New $16,000 Postpartum Depression Drug Is Here. How Will Insurers Handle It?

Case study sources: Veozah (fezolinetant)

- FDA. (2023). Veozah (fezolinetant) Prescribing Information.
- Astellas Pharma. (2024). VEOZAH Access and Savings. Manufacturer's website.
- Elendu et al. (2024). FDA Approves Veozah (Fezolinetant) for Menopausal Symptoms: A New Nonhormonal Option. Annals of Medicine and Surgery.

Case study sources: Orilissa (elagolix)

- FDA. (2018). Orilissa (elagolix) Prescribing Information.
- GoodRx. (2025). How Much Does Orilissa Cost Without Insurance?
- AbbVie. Orilissa Patient Support and Cost Information.

Case study sources: Makena (17-OHPC)

- FDA. (2023 April 5). FDA Commissioner and Chief Scientist Announce Decision to Withdraw Approval of Makena.
- Nelson, D.B. et al. (2023). The End Is Where We Start From: Withdrawal of 17-Alpha Hydroxyprogesterone Caproate. American Journal of Obstetrics & Gynecology.
- Society for Maternal-Fetal Medicine. (2023). Statement on FDA Withdrawal of 17-OHPC. American Journal of Obstetrics & Gynecology.
- NPR Health. (2020 January 24). Drug To Prevent Premature Birth Divides Doctors, Insurers And FDA Experts.

Case study sources: Renovia Leva

- Weinstein MM et al. (2022). Digital Therapeutic Device for Urinary Incontinence: A Randomized Controlled Trial. Obstetrics & Gynecology/PMC.
- Weinstein, M.M. et al. (2023). Digital Therapeutic Device for Urinary Incontinence: A Longitudinal Analysis at 6 and 12 Months. Obstetrics & Gynecology.
- ObvioHealth. (2022). ObvioHealth Completes First Fully Virtual Urogynecology Clinical Trial.
- ObvioHealth. DCTs and Devices: Lessons Learned from a Breakthrough Medical Device Trial in Women's Health.
- Renovia. (2022). Renovia Receives 2021 Excellence Award for leva.

Case study sources: Entresto (sacubitril-valsartan)

- Novartis. (2021 February 16). Novartis Entresto Granted Expanded Indication in Chronic Heart Failure by FDA.
- The Medical Letter. (2021). In Brief: Sacubitril/Valsartan (Entresto) Expanded Indication.
- Solomon S.D. et al. (2021). Potential Implications of Expanded US Food and Drug Administration Labeling for Sacubitril/Valsartan in the US. JAMA Cardiology/Medscape.

Policy and regulatory sources

- FDA. (2024). Diversity Action Plans to Improve Enrollment of Participants from Underrepresented Populations in Clinical Studies: Guidance for Industry.
- FDA. (2024). Conducting Clinical Trials With Decentralized Elements.
- ICH. (2023). ICH E21: Inclusion of Pregnant and Breastfeeding Individuals in Clinical Trials.
- NIH Office of Research on Women's Health. (2016, ongoing). Sex as a Biological Variable.

About ZS

ZS is a management consulting and technology firm that partners with companies to improve life and how we live it. We transform ideas into impact by bringing together data, science, technology and human ingenuity to deliver better outcomes for all. Founded in 1983, ZS has more than 15,000 employees in over 40 offices worldwide. To learn more, visit www.zs.com or follow us on LinkedIn.