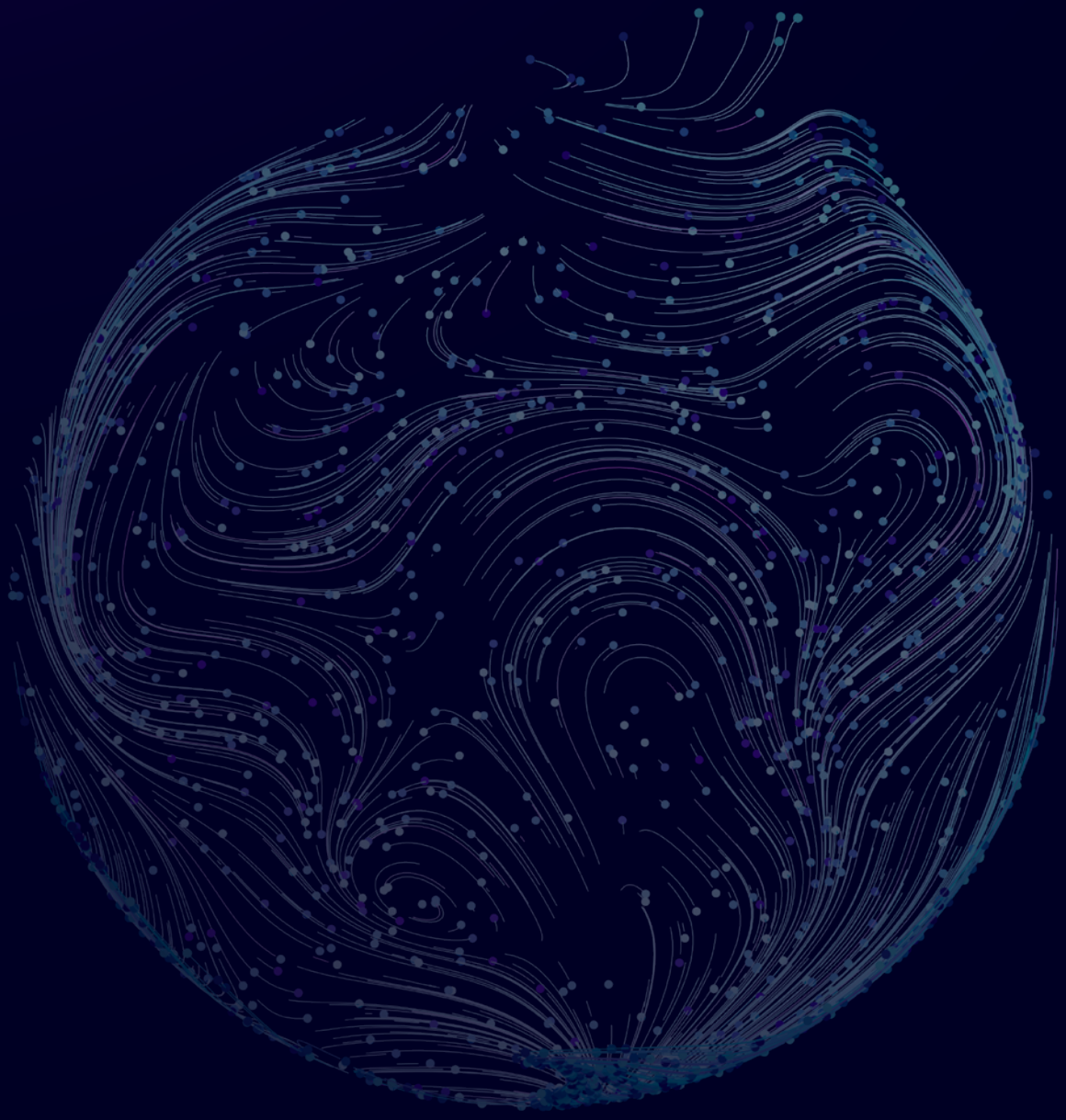
Scientific Innovations
from 2025

Citation: EMJ Innov. 2026;10[1]:50-61.
https://doi.org/10.33590/emjinnov/TTUK8215

The following highlights showcase some of the most significant, innovative research published across multiple therapeutic areas in medicine in 2025. The selected abstracts mark key advancements in healthcare and highlight future avenues for medical innovation.



Automated Insulin Delivery Boosts Type 2 Diabetes Control

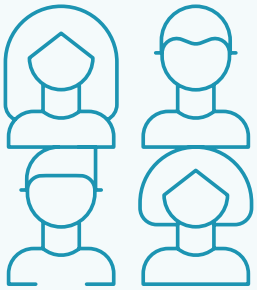
A GROUNDBREAKING multicentre trial has shown that automated insulin delivery (AID) systems can substantially improve glucose control in adults with insulin-treated Type 2 diabetes. While AID has previously proven effective in Type 1 diabetes, high-quality evidence for its use in Type 2 diabetes has been lacking.[1]

The 13-week randomised study enrolled 319 participants who were assigned in a 2:1 ratio to either an AID system or to continue their usual insulin regimen, with all participants using continuous glucose monitoring. The trial's primary outcome was the change in HbA1c at Week 13.

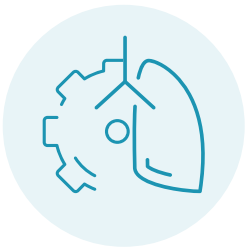
Findings revealed a marked improvement in glycaemic control for those using AID. HbA1c decreased from 8.2% to 7.3% in the AID group, compared with a modest reduction from 8.1% to 7.7% in the control group, representing a significant mean difference of -0.6 percentage points. Time spent within the target glucose range (70–180 mg/dL) increased from 48% to 64% with AID, while the control group showed virtually no change. Secondary measures of hyperglycaemia also favoured AID, highlighting its ability to maintain glucose levels closer to optimal targets. Hypoglycaemia was rare across both groups, with only a single severe event reported in the AID arm.

The results suggest that AID systems can deliver clinically meaningful improvements over standard insulin therapy with continuous glucose monitoring alone. By automating insulin adjustments based on real-time glucose readings, these systems reduce the burden on patients while enhancing safety and efficacy.

As diabetes management increasingly embraces technology, AID represents a significant step towards personalised, responsive care for people with Type 2 diabetes. The trial provides compelling evidence that automated systems, once primarily reserved for Type 1 diabetes, may soon play a pivotal role in optimising outcomes for a broader patient population.



The 13-week randomised study enrolled 319 participants who were assigned in a 2:1 ratio to either an AID system or to continue their usual insulin regimen



Alzheimer's Disease Prediction: A Multi-agent Large Language Model Framework

A MULTI-AGENT large language model framework has demonstrated that it can feasibly support early risk assessment for Alzheimer's disease by extracting clinically relevant indicators from unstructured records and generating more accurate long horizon predictions than single model approaches.[2]

Early detection of Alzheimer's disease relies heavily on assessments and biomarker tests that are expensive, invasive, or impractical for population-level use. Although structured electronic health record data have been used for risk prediction, unstructured clinical narratives contain valuable information that is often overlooked. Recent advances in language models have created opportunities to analyse free text at scale; however, challenges persist regarding privacy, scalability, and the limited reasoning capacity of single models for complex diagnostic tasks. These limitations have prompted interest in collaborative language model systems that emulate multidisciplinary clinical reasoning.

To address these issues, researchers developed CARE-AD, a multi-agent framework for forecasting Alzheimer's disease onset using longitudinal electronic health record notes. The system assigns specialised agents to clinical domains including primary care, neurology, psychiatry, geriatrics, and psychology. Each agent reviews temporally ordered narratives and produces a domain-specific assessment of symptoms and functional decline. A separate Alzheimer's specialist agent integrates these assessments into a single risk estimate. The architecture is designed to capture early and often subtle patterns that may be underrepresented in structured data, while providing interpretable, agen-level outputs that clinicians can examine.

In a retrospective study, using data from the United States Veterans Health Administration (VHA), CARE-AD achieved higher accuracy than baseline single model methods across all prediction windows. At 10 years prior to the first diagnosis code, accuracy was 0.53, compared with 0.26–0.45 for benchmark approaches. Performance remained strong at closer intervals, reaching 0.83 at 1 day before diagnosis. These findings show that clinically meaningful signals appear in unstructured narratives well before formal diagnosis, and that multi-agent modelling improves detection of such patterns.

The results indicate that multi-agent language model systems may offer a promising pathway for earlier identification of Alzheimer's disease and could enhance clinical decision support by providing transparent, domain-specific reasoning. Future work should include prospective validation, assessment across additional populations, and integration with workflow-compatible clinical tools to determine how such systems can best support risk stratification and targeted early intervention in practice.

> **"The architecture is designed to capture early and often subtle patterns that may be underrepresented in structured data"**

Digital Self-Management Programmes Enhance Asthma Control

DIGITAL health tools are increasingly explored as scalable strategies to support asthma self-management, yet robust evidence for their long-term effectiveness remains limited. In this pragmatic randomised clinical trial, investigators evaluated whether a digital asthma self-management (DASM) programme could improve symptom control among adults with asthma compared with usual care.[3]

The decentralised, open-label study enrolled 901 adults with asthma, who were recruited remotely and followed for 12 months, reflecting real-world implementation outside of traditional clinical settings.

Participants randomised to the DASM intervention used an app-based platform incorporating symptom logging, tailored notifications, wearable device integration, and self-management tools, while control participants received usual care. The primary outcome was change in Asthma Control Test (ACT) score, a validated patient-reported measure of asthma control. Secondary outcomes included engagement with the digital platform and self-reported medication adherence.

Among participants with uncontrolled asthma at baseline, the DASM programme was associated with significantly greater improvement in asthma control compared with usual care over 12 months. DASM users experienced a mean ACT improvement of 4.6 points, compared with 1.8 points in the control group, resulting in a clinically and statistically significant adjusted between-group difference. These findings suggest that digital self-management support can meaningfully enhance symptom control when added to standard care.

Subgroup analyses revealed that race moderated the observed treatment effect. While non-African American participants demonstrated substantial benefit from DASM, the difference between intervention and control groups among African American participants was smaller and not statistically significant. No significant moderation was observed by insurance status or Hispanic ethnicity, indicating broadly similar effects across these sociodemographic factors. The reasons underlying differential response by race remain unclear and warrant further investigation.

Overall, this trial provides evidence that a DASM programme can improve long-term symptom control in adults with asthma, particularly among those with uncontrolled disease at baseline. The findings support continued development and refinement of digital asthma interventions, with attention to ensuring equitable effectiveness across diverse populations.

H-Scan Ultrasound Provides Rapid, Accurate Kidney Fibrosis Assessment

IN A NEW study, researchers have developed a quantitative algorithm, called renal H-scan, that accurately estimates kidney quality at the time of considering it as a donor organ. This ensures that donor recipients have the highest chance of survival by receiving a fully functional, high-quality kidney.[4]

H-scan estimates of whole-kidney fibrosis were found to correlate closely with renal function following transplantation

Kidney transplantation is currently the optimal treatment for renal failure. The current procurement process in the USA involves carrying out a biopsy to measure renal fibrotic burden, a critical measure of irreversible kidney injury. However, these biopsies come with limitations. Inaccuracies can be introduced by sampling bias and rapid sample preparation. Moreover, biopsy carries a risk of bleeding and requires continuous access to trained pathology expertise, which is not always feasible. Importantly, the highly localised nature of biopsy samples often fails to represent fibrosis across the whole kidney, limiting their predictive value for post-transplant renal function.

In a first-in-human study, researchers have developed renal H-scans, which can be integrated into standard ultrasound workflows as another option to measure the fibrotic burden quickly and non-invasively. The H-scan algorithm has been tested in preclinical animal models and human transplant kidneys, offering the ability to assess fibrosis across the entire organ rather than a small, localised sample. Unlike traditional biopsy, H-scan estimates of whole-kidney fibrosis were found to correlate closely with renal function following transplantation, suggesting superior predictive potential for post-transplant outcomes.

This innovation has several important implications. By providing an accurate, non-invasive assessment of kidney quality, H-scans could reduce reliance on biopsies, minimising procedural risks and the need for round-the-clock pathology support. The technique is straightforward to implement within existing ultrasound workflows, making it accessible and scalable for clinical practice. Additionally, accurate whole-kidney fibrosis quantification could improve organ allocation decisions and help ensure that transplant recipients receive kidneys with the highest likelihood of supporting long-term function.

Overall, the renal H-scan represents a novel, practical, and clinically impactful tool for improving the assessment of donor kidney quality. Its adoption could enhance patient safety, optimise transplant outcomes, and address critical limitations associated with biopsy-based fibrosis assessment.

Deployment of a Digital Clinical Alert Navigation System in Integrated Care

A RESEARCH study has demonstrated that natural language processing-driven clinical alerts and navigation tools can improve high acuity detection, support appropriate care routing, and contribute to an enhanced patient experience within an integrated, value-based care model.[5]

Patient portals have become central to digital self-service in healthcare, yet ensuring that such systems reliably guide patients towards clinically appropriate pathways remains a major challenge. In response to growing concerns related to self-scheduling, the Southern California Permanente Medical Group (Pasadena, USA) established the Virtual Safety Net, a natural language processing-enabled ecosystem that successfully identified time sensitive risks, but could not be readily scaled. These results highlighted the need for a more comprehensive, adaptable system, capable of embedding safety, navigation accuracy, and user centric design into routine digital interactions. Therefore, the Kaiser Permanente Intelligent Navigator (KPIN; Kaiser Permanente, Oakland, California, USA) was developed to integrate clinical alerts and navigation tools into the appointment booking workflow.

The KPIN system processes input through a multilingual natural language pipeline that incorporates large language models and custom transformer architectures. These models evaluate symptoms, identify high acuity presentations, retrieve relevant demographic and clinical guideline data, and generate a curated set of appropriate care options. Outputs include video or phone visits, asynchronous messaging, or other modalities consistent with clinical standards. Encounters conclude when patients select an offering, whereas abandonment represents incomplete interaction.

KPIN demonstrated strong performance in detecting high acuity symptoms, achieving 96.0% accuracy (95% CI: 93.7–98.0%), with 97.5% precision (95% CI: 95.8–99.0%), and a recall of 96.0% (95% CI: 93.8–97.9%). Similarly, the clinical navigation models achieved 81.9% accuracy (95% CI: 80.0–83.6%), with a corresponding precision of 85.6% (95% CI: 84.0–87.2%), a recall of 81.9% (95% CI: 80.1–83.7%), and an F1-Score of 82.8% (95% CI: 81.1–84.5%). Additional metrics showed that KPIN's adjusted successful booking rate was 53.68%, with an abandonment rate of 2.94% (interquartile range: 2.77–3.11%), aligning with patient survey results showing an 8.63 percentage point increase for positive sentiment.

These findings indicate that an integrated digital navigation system can enhance clinical safety, streamline access, and support value-based care by directing patients to suitable modalities with greater accuracy. Future work should focus on establishing pre-implementation baselines, expanding conversational capabilities, and evaluating real-world clinical outcomes to determine how such systems can best complement clinician workflows and improve timely access to appropriate care.

New Microneedle Device Tracks Lactate in Pregnancy

RESEARCHERS in Liverpool, UK, have successfully tested a new microneedle device that can continuously monitor lactate levels in pregnant women, a development that could transform maternity care by enabling earlier detection of complications such as sepsis during labour.[6]

The pilot study, conducted at a clinical research facility in a Liverpool hospital, involved seven healthy, pregnant volunteers over the age of 18 years. All participants had uncomplicated pregnancies and were accustomed to light exercise. Each woman wore a small microneedle patch designed to measure lactate levels in the skin's interstitial fluid, while simultaneous blood samples were taken to compare results.

During a 30-minute session of gentle exercise followed by rest, the device tracked changes in lactate levels continuously. Researchers found that interstitial lactate closely mirrored venous lactate trends, suggesting that the new technology could offer a reliable, non-invasive alternative to traditional blood testing. Participants reported minimal discomfort, with average pain and discomfort scores of just 0.43 and 0.14 out of 10, respectively.

As stressed by the researchers, little is known about lactate levels during labour, with some expectation of them rising due to anaerobic respiration during physical exertion. Elevated lactate levels, particularly above 2 mM, can prompt clinicians to consider sepsis, a potentially life-threatening condition requiring immediate treatment. Current lactate measurements provide only snapshots in time, but the new device offers continuous, real-time insight into a patient's condition.

Although this proof-of-concept study involved only a small group of healthy volunteers, researchers say the findings point to significant future applications. Continuous lactate monitoring during labour could help clinicians identify early signs of distress or infection, enabling swift intervention while also avoiding unnecessary use of antibiotics.

Continuous lactate monitoring during labour could help clinicians identify early signs of distress or infection, enabling swift intervention while also avoiding unnecessary use of antibiotics

Virtual Reality Meditation Eases Depression and Anxiety

A NEW clinical study has found that immersive virtual reality meditation (IVRM) may offer a promising, non-pharmacological option for reducing symptoms of major depressive disorder and generalised anxiety disorder among hospital inpatients. The research adds to growing interest in technology-enabled mindfulness therapies, which aim to increase access to evidence-based psychological support.[7]

Mindfulness-based cognitive therapy has long been recognised as an effective complement to traditional treatment for depression and anxiety. However, questions remain about whether technology-assisted approaches can achieve similar clinical benefits, particularly within inpatient settings where symptom severity is often higher. This study explored whether IVRM, an immersive adaptation of mindfulness-based cognitive therapy, could improve emotional regulation and, in turn, reduce depressive and anxiety symptoms.

The 10-week, single-arm clinical trial involved 26 participants at a community hospital behavioural health unit, each diagnosed with major depressive disorder and generalised anxiety disorder. Patients engaged in IVRM sessions three times per week, with outcomes monitored using depression and anxiety assessments alongside ECG-based measures of emotional regulation. The study focused particularly on the Coherence Achievement Score (CAS), a physiological indicator reflecting the stability and balance of emotional responses.

Researchers analysed associations between CAS, symptom changes, and relevant covariates using a generalised estimating equation model. Results showed a clear relationship: improvements in emotional regulation following IVRM sessions were associated with meaningful reductions in both depression and anxiety. The findings support the idea that virtual reality may not only enhance engagement with mindfulness practices, but also strengthen the emotional regulatory mechanisms underpinning therapeutic improvement.

While the study is limited by its small sample size and single-arm design, it contributes valuable insights into the feasibility and clinical potential of IVRM in inpatient mental health care. As interest grows in innovative, accessible mental-health interventions, immersive virtual reality may become an important tool for hospitals seeking complementary treatments that reduce symptom burden without adding pharmacological side-effects.

Carbon Emissions of Virtual Wards Compared with Inpatient Care

DIGITALLY enabled models of care are increasingly central to healthcare delivery within the NHS, offering opportunities to expand capacity while supporting sustainability goals. As virtual wards are scaled nationally, understanding their environmental impact is essential to aligning innovation with the NHS commitment to achieving net zero carbon emissions. This study evaluated the carbon footprint of a virtual ward pathway for patients with acute respiratory infections and frailty in a large acute hospital trust, using a retrospective cohort analysis.[8]

Carbon emissions associated with virtual ward care were compared with those generated by traditional inpatient bed days. The analysis demonstrated that virtual wards were associated with a substantially lower carbon impact, largely driven by reductions in hospital-based resource use and patient travel. These findings support the premise that home-based digital care pathways can contribute meaningfully to healthcare decarbonisation when implemented at scale.

The study also compared two approaches to carbon accounting: a manual data audit and an automated business intelligence extraction. While automated methods offered efficiency and scalability, manual audits provided more granular insights into emissions related to community care and out-of-hospital activity. This discrepancy highlights current limitations in routinely collected datasets when assessing the full environmental impact of digital care models.

Overall, the findings suggest that virtual wards represent a lower-carbon alternative to inpatient care for selected patient groups. However, accurate measurements of their environmental benefit depend on improved data capture beyond the hospital setting. Enhancing automated systems to better reflect community-based care will be critical for robust carbon evaluation as digitally enabled services continue to expand across the NHS.

> **The analysis demonstrated that virtual wards were associated with a substantially lower carbon impact, largely driven by reductions in hospital-based resource use and patient travel**

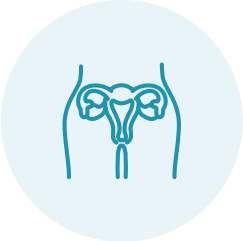
Wearable Sweat Sensor Enables Remote Monitoring in Cystic Fibrosis

SWEAT chloride concentration is a key biomarker in cystic fibrosis (CF), traditionally measured using pilocarpine-induced sweat collection and chloridometry. While this method remains the diagnostic gold standard, it is impractical for frequent or remote monitoring. This study evaluated a wearable skin-interfaced microfluidic 'CF Patch' (Epicore Biosystems, Cambridge, Massachusetts, USA), paired with smartphone-based image analysis, as a tool for real-time and remote assessment of sweat biomarkers in adults with CF.[9]

Clinical studies compared CF Patch measurements with standard chloridometry under both pilocarpine-induced and exercise-induced sweating conditions. In laboratory settings, the CF Patch demonstrated strong correlations with sweat chloride values obtained via pilocarpine-based chloridometry, regardless of whether sweat was induced pharmacologically or through exercise. These findings support the accuracy of the device for quantifying sweat chloride and sweat rate in controlled environments.

The feasibility of remote monitoring was also assessed. In healthy volunteers, exercise-induced sweat chloride measurements collected remotely using the CF Patch were strongly correlated with in-laboratory CF Patch measurements. In contrast, correlations were weaker in people with CF, particularly among those receiving cystic fibrosis transmembrane conductance regulator (CFTR) modulator therapy. Importantly, individuals with CF on modulators showed greater day-to-day variability in sweat chloride levels compared with healthy volunteers. This variability highlights a limitation of relying on single, in-laboratory chloridometry measurements to assess CFTR modulator efficacy and pharmacodynamics.

Overall, the findings indicate that the CF Patch enables the serial, non-invasive measurement of sweat chloride in both laboratory and remote settings. While the device is not intended to replace pilocarpine-induced chloridometry for diagnostic purposes, it shows promise as a remote disease management tool. By allowing repeated measurements over time, the CF Patch may provide more nuanced insights into treatment response and medication effectiveness in adults with CF.



Leveraging Large Language Models to Personalise Therapy in Rare Gynaecologic Malignancies

RARE gynaecological tumours (RGT) remain a difficult concept in oncology due to their low incidence, pronounced heterogeneity, and absence of robust, evidence-based clinical guidelines. These challenges often result in delayed diagnosis, limited therapeutic direction, and poor clinical outcomes for affected patients. While molecular tumour boards provide a promising route towards individualised therapy by leveraging biomarker-guided decision-making, their impact is frequently constrained by the fragmented, unstructured nature of clinical and molecular data. Manual data curation slows the process, limits scalability, and increases the risk of overlooking relevant therapeutic opportunities.[10]

The authors developed a proof-of-concept LLM-enabled digital twin framework combining data from institutional cases and published case reports (21 patients), with a comprehensive literature dataset derived from 655 peer-reviewed publications

This study evaluates whether large language models (LLM) can overcome these barriers by enabling the construction of digital twins, computational patient replicas that integrate detailed clinical, pathological, and biomarker information. The authors developed a proof-of-concept LLM-enabled digital twin framework combining data from institutional cases and published case reports (21 patients), with a comprehensive literature dataset derived from 655 peer-reviewed publications. The system was applied to metastatic uterine carcinosarcoma, a highly aggressive and understudied RGT with limited treatment consensus.

The LLM-powered digital twins demonstrated the capability to synthesise multimodal data streams, harmonise terminology, and generate personalised treatment recommendations. Importantly, the digital twin approach identified potential therapeutic strategies not captured through traditional, single-source analyses, highlighting the value of cross-referencing diverse datasets and expanding the evidence base for rare, complex cancers. By modelling possible clinical trajectories, the system also offers a means to anticipate disease progression and refine therapy choices over time.

This work highlights a significant conceptual shift in oncology: moving from an organ-based categorisation of tumours to a biology-driven, biomarker-centric framework. LLM-enabled digital twins have the potential to accelerate precision oncology for RGTs by improving data accessibility, reducing manual workflow burden, and more effectively linking patients to targeted therapies. As digital twin technologies mature, they may help clinical teams overcome the inherent evidence gaps in rare tumour research, ultimately supporting more individualised care pathways and improving outcomes for patients with rare gynaecological malignancies.

References

1. Kudva YC et al.; 2IQP Study Group. A randomized trial of automated insulin delivery in type 2 diabetes. N Engl J Med. 2025;392(18):1801-12.
2. Li R et al. CARE-AD: a multi-agent large language model framework for Alzheimer's disease prediction using longitudinal clinical notes. NPJ Digit Med. 2025;8(1):541.
3. Silberman J et al. a digital asthma self-management program for adults: a randomized clinical trial. JAMA Netw Open. 2025;8(7):e2521438.
4. Hysi E et al. A first-in-human study of quantitative ultrasound to assess transplant kidney fibrosis. Nat Med. 2025;31:970-8.
5. Nguyen D et al. Digital transformation with clinical alerts and personalized care systems in an integrated value based model. NPJ Digit Med. 2025;8:415.
6. Freeman DME et al. First in pregnancy continuous lactate monitoring using a minimally invasive microneedle device: a proof of concept study in healthy pregnant volunteers. BMJ Innovations. 2025;DOI:10.1136/bmjinnov-2025-001393.
7. Lee J et al. The impact of immersive virtual reality meditation for depression and anxiety among inpatients with major depressive and generalized anxiety disorders. Front Psychol. 2024;15:1471269.
8. Townsend K et al. Exploring the carbon impact of virtual wards in a large acute hospital. BMJ Innovations. 2025;11:19-125.
9. Nelson RS et al. Remote analysis and management of sweat biomarkers using a wearable microfluidic sticker in adult cystic fibrosis patients. Proc Natl Acad Sci U S A. 2025;122(33):e2506137122.
10. Lammert J et al. Large language models-enabled digital twins for precision medicine in rare gynecological tumors. NPJ Digit Med. 2025;8:420.