



Bridging Innovation to Implementation in AI Fracture Detection: A Commentary Piece

Journal:	<i>The Bone & Joint Journal</i>
Manuscript ID	BJJ-2024-1567.R1
Manuscript Type:	Other
Keywords:	Artificial Intelligence, Implementation, Imaging

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Bridging Innovation to Implementation in AI Fracture Detection: A Commentary Piece

Introduction

We stand at the face of a revolution in healthcare, with the advance in Artificial intelligence (AI). Governments worldwide are actively supporting AI development, aiming to enhance efficiency and optimise healthcare delivery.¹ The Darzi Report (2024) emphasised the need for a “major tilt” towards technology to unlock productivity and the potential of AI to transform healthcare delivery.² However, traditionally, there is a staggering 17-year lag between the generation of evidence and its integration into clinical practice.³ Can we speed up and improve the adoption by understanding implementation pathways?

The application of AI seems particularly applicable in the realm of medical imaging, through the ability to transform orthopaedic care. This commentary explores the evolution of AI-driven fracture detection X-ray (XR) systems, emphasising their value proposition and the need to bridge the gap between technological advancements and their practical application within clinical settings. In particular we explore the implementation challenges that hinder its adoption. There is a slowly growing effort to identify these challenges and a shift towards the development of strategies to optimise implementation.⁴

AI-driven fracture detection

Traditional methods of fracture detection rely on manual interpretation of XR images, which can be subjective, time-consuming, and prone to variability among clinicians. AI-based fracture detection systems offer a promising solution to overcome these challenges. These algorithms utilise machine learning techniques, typically in the form of deep learning models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), to analyse X-ray images, identify fracture patterns, and differentiate them from normal anatomical structures. By leveraging the power of AI, these algorithms can process vast amounts of medical data, learn from patterns, and make accurate predictions and diagnostic tasks.⁵ The ability of AI algorithms to analyse images quickly and accurately has the potential to enhance diagnostic accuracy, reduce interpretation variability, and optimise workflow efficiency in fracture detection.⁶

Recent evidence demonstrates that AI performs comparably to human radiologists in detecting wrist, ankle, and hand fractures, while outperforming them in identifying hip and proximal humerus fractures. Pooled sensitivity and specificity in these meta-analyses consistently exceed 90%, underscoring AI’s potential to enhance diagnostic accuracy and reduce errors.^{7, 8}

AI's clinical value is further highlighted in individual studies. Olczak et al. (2017)⁹ demonstrated AI's effectiveness in orthopaedic trauma radiographs, matching senior surgeons in accuracy while aiding junior clinician decision-making. Urakawa et al. (2018)¹⁰ showed that deep learning models detected intertrochanteric hip fractures with an AUC of 0.98, achieving equivalent performance to expert orthopaedic surgeons while improving workflow efficiency. Chung et al. (2018)¹¹ found AI outperformed radiologists in detecting proximal humerus fractures, reducing inter-reader variability.

Despite these promising findings, real-world implementation remains limited. Langerhuizen et al. (2019)¹² emphasised that while AI demonstrates high sensitivity, specificity challenges remain, particularly in cases where normal anatomical variations mimic fractures. The need for multi-centre validation and external testing is critical to ensure generalizability across diverse patient populations.

The Problem

A recent state-of-the-art review discussed the available AI fracture detection systems, ready for integration into radiological workflows for both adult and paediatric patients. They found 21 commercially-available AI products with Conformité Européenne (CE)/Food and Drug Administration (FDA) approval, however there is particularly slow uptake across the 215 UK hospital trusts.¹³

What is the Value Proposition?

1. Bimodal Fracture Incidence Peaks

Fractures represent a significant burden on healthcare systems, with peaks in the paediatric and elderly populations.¹⁴ At their childhood peak, the incidence of fractures is only ever surpassed by osteoporotic females at 85 years of age and never among men.¹⁵ There are distinctive challenges posed by fractures in childhood which necessitate innovative solutions to streamline diagnosis and treatment. The deployment of AI-driven fracture detection tools introduces a transformative paradigm in fracture diagnostics, especially in these cohorts.

The key challenge in children is fractures involving growth plates. These comprise up to 18% of all paediatric fractures.¹⁶ The intricate and sometimes very subtle nature of growth plate fractures can mislead even the most experienced emergency clinicians or orthopaedic surgeons, leading to misdiagnoses. The consequences of misdiagnoses in this vulnerable population can be profound, through delayed intervention or abnormal growth. This can lead to significant deformity and pain, sometimes with debilitating effects on patient quality of life which may necessitate corrective surgery. Whilst injuries may be missed, perhaps more commonly growth plates are misinterpreted as fractures, especially in the elbow and ankle, which are areas where the appearance of the growth plates are

particularly challenging.^{17,18} The overdiagnosis leads to unnecessary treatment (i.e. immobilisation) for the child, unnecessary anxiety and inconvenience for the family and unnecessary appointments in fracture clinics.

In the elderly population, delayed detection and intervention significantly contribute to adverse outcomes. Protracted waiting times in the emergency department prior to referral to orthopaedics is an important bottleneck, which can be exacerbated by fractures which may necessitate more senior advice or further imaging to definitively characterise fractures.¹⁹ This is commonly seen in fragility hip fractures, where it is routine practice to arrange a computed tomography (CT Scan) to confirm the diagnosis in equivocal cases.²⁰ Early definitive care is lost during these delays, exacerbating pain and increasing risk of significant long-term morbidity and mortality to this vulnerable patient group. There are promising results in the potential applications of AI to aid diagnosis of hip fractures on radiographs.²¹

2. Scale of Financial Implications on the NHS

Unnecessary Appointments

The average cost per patient appointment, within the traditional fracture clinic pathway, is £36.81. This is important for evaluating the economic implications of inappropriate referrals.²² According to “Department of Health Audit Commission. Outpatients – Review of National Findings” from the Department of Health (2003), 75-92% of all musculoskeletal (MSK) injuries in the NHS trigger follow-up with Orthopaedic surgeons, leading to 7.1 million appointments annually. Among these, 30% account for new cases, while 70% constitute follow-up reviews; this means around 2.13 million new patient referrals. Notably, one in six new patients is discharged at first follow-up.²³ Furthermore, it is suggested that 23% of cases can be discharged directly from the emergency department (ED) with the introduction of new, more efficient pathways, avoiding unnecessary orthopaedic referrals.²⁴ This translates to almost 500,000 unwarranted orthopaedic appointments for new patients with MSK injuries and an annual NHS cost of £18 million.

These findings underscore the need for an effective tool to streamline fracture diagnosis and facilitate better decision making. The potential for substantial gains in health, alongside cost-savings to the NHS is clear.

Misdiagnosis

An analysis of diagnostic errors in a single emergency department reported that 79.7% were attributed to missed fractures, the overwhelming majority of which were as a result of misreading radiographs,

this is particularly common in undisplaced fractures being scrutinised by an inexperienced workforce; this highlights the clear value in bridging this skill-gap.^{25, 26}

In the 2023/24 NHS Resolution Annual Report, the total value of clinical claims amounted to £6.5778 billion. Of this, Orthopaedics accounted for 3.2% (£210.49 million), and Emergency Medicine contributed 7.1% (£467 million). An estimated 5.7% of all Emergency Department (ED) visits involve a diagnostic error, and 79.7% of those errors are attributed to missed fractures, largely due to the misinterpretation of radiographs. By applying these figures, it is estimated that over £20 million of the total Emergency Medicine claims can be attributed to missed fractures, underscoring the substantial financial burden these errors place on the NHS.²⁷

By implementing AI-driven fracture detection systems, there is a significant opportunity to mitigate these diagnostic errors, thereby reducing litigation costs and improving both patient outcomes and resource utilisation within the NHS.

Health Economic Evaluations and Cost-Effectiveness

Despite the potential economic benefits of AI implementation, financial constraints—particularly within the NHS—remain a significant barrier to adoption. The initial investment required for AI integration, including software acquisition, infrastructure development, and staff training, is often substantial. While AI has the potential to reduce healthcare costs by improving efficiency and reducing diagnostic errors, real-world cost-effectiveness analyses are lacking. Voets et al. (2022) highlight that health economic evaluations (HEEs) of AI in healthcare are limited and predominantly focus on costs rather than health outcomes, in their systematic review. Notably, both short-term and long-term model-based evaluations are rare, resulting in a lag in understanding AI's actual benefits relative to technological advancements. Future research should prioritise comprehensive economic evaluations that quantify AI's return on investment, providing clear evidence of its long-term financial sustainability and justifying its widespread adoption.²⁸

3. Resource Utilisation Optimisation:

Accelerated diagnostic process directly correlates with resource optimisation. EDs often contend with limited resources, including healthcare personnel and equipment. The more efficiently these resources are utilised, the more patients can be attended to without compromising the quality of care. AI's ability to swiftly and accurately identify fractures ensures that each resource is employed judiciously, minimising waste and maximising impact. The optimised use of medical personnel and equipment prevents unnecessary expenditures, aligning with the NHS's broader goals of fiscal responsibility. The innovative nature of this frontier positions it as a catalyst for transformative change

in AI medical imaging solutions and offers a multi-faceted benefit, from individual patient care to the broader economic considerations of healthcare delivery within the NHS.²⁹

Broader Applications of Machine Vision in Orthopaedic Surgery

While AI-driven fracture detection holds significant promise, other AI applications within orthopaedic imaging—collectively known as machine vision—are also emerging. These tools, some already in use and others on the brink of routine adoption, include skeletal age assessment in children³⁰, early detection of hip disease in cerebral palsy³¹, and identifying signs of implant failure following joint replacement surgery³². Exploring the implementation of multiple AI tools offers a broader perspective on their integration into clinical practice, helping to identify common barriers, facilitators, and implementation strategies. By examining these diverse use cases, we can gain valuable insights into the wider adoption of AI in orthopaedic surgery beyond fracture detection. and therefore formulate a comprehensive blueprint for its effective implementation.

What are the Challenges to AI Medical Imaging Adoption?

Whilst AI has demonstrated its value in improving diagnostic accuracy, its integration into routine clinical workflows presents substantial challenges, contributing to its slow uptake. These can be considered barriers to implementation and include factors such as regulatory approvals, clinician acceptance, ethical considerations, and financial constraints.

One of the most significant barriers to AI adoption is trust—both among clinicians and patients. According to Ormond et al. (2023), there remains widespread scepticism regarding AI’s role in medical decision-making, with concerns surrounding autonomy, reliability, and medico-legal implications. Clinicians often perceive AI as a ‘black box’ technology, where decision-making processes are opaque and difficult to interpret. Addressing these concerns requires the development of explainable AI (XAI) models, which provide transparent and interpretable outputs, allowing clinicians to understand and trust AI-generated recommendations.³³

Moreover, AI models are susceptible to algorithmic biases, particularly if trained on homogeneous datasets that do not reflect real-world patient demographics. This raises concerns about health disparities, where AI models may perform sub-optimally in underrepresented populations.

Importance of Implementation Science Theory

Implementation science is a field of study that examines the methods and strategies required to promote the systematic uptake of research findings and evidence-based practices. While technological advancements, particularly in AI, are rapidly transforming healthcare, the effective

assimilation of these innovations into clinical practice remains a challenge. Implementation Science provides a structured framework to systematically study, understand, and address the wide array of factors influencing the successful integration of AI solutions in healthcare. The potential exists to gain insight, with the aim of improving and expediting the integration of AI technology into healthcare.³⁴ Classically this is done retrospectively in cases of implementation failure, however there is a drive to develop implementation strategies with prospective analysis to optimise deployment.³⁵ Nilsen et al. have described a clear need for the development of an AI Quality Implementation Framework (AI-QIF) as a tool to guide the implementation of AI-based applications in health care, however no such tool exists.³⁶

Determinant frameworks, such as the Consolidated Framework for Implementation Research (CFIR), are a subset of frameworks within the speciality, that provide a theoretical lens, employing 5 core domains, for in-depth exploration of the intricate dynamics between technology, clinicians, patients, and organisational systems, to identify barriers and facilitators that influence implementation. Normalisation Process Theory (NPT) is a specific determinant model that helps us understand the adoption, implementation and sustainment of complex interventions or innovations, ultimately driving integration into routine practice. Evaluation frameworks, such as RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance), offer insights into the success and outcomes of the implementation strategy used, enabling further refinement, optimisation and improvement. By applying such implementation science principles, there is potential to transcend the conventional boundaries of technological research, aiming to not just scrutinise innovative AI-driven systems but to strategically navigate the intricate landscape of healthcare delivery and provide a blueprint for successful deployment. The goal is to ensure that the benefits of AI in medical imaging solutions are not just theoretical but tangibly realised in the daily practices of healthcare professionals, ultimately enhancing patient care and reshaping services within the NHS.^{35,37,38}

To optimise AI implementation, Farrow et al. (2024) introduced the CPI-AI framework, integrating AI deployment within the IDEAL principles (Idea, Development, Exploration, Assessment, Long-term study), providing a structured pathway for evaluating AI implementation at different stages. This framework ensures that AI solutions undergo rigorous validation before being deployed into high-stakes clinical environments.³⁹

The Knowledge-to-Action (KTA) Cycle is a landmark model within implementation science that bridges the gap between research knowledge and practical application. Developed by Graham et al., this theory outlines the process of translating scientific knowledge into actionable healthcare practices, policies, and interventions. The KTA Cycle consists of two main components: Knowledge Creation and

the Action Cycle. We have adapted this concept to streamline its prospective use in the deployment of novel technology (Figure 1).⁴⁰

Concluding Remarks

As we contemplate the future of healthcare, the integration of artificial intelligence, especially in orthopaedic medical imaging, presents a paradigm shift with immense potential. However, the real challenge lies in translating these technological advancements into everyday clinical practice. While AI has shown promising capabilities in enhancing diagnostic accuracy and streamlining workflows, a significant gap persists between its theoretical potential and practical application. Bridging this gap necessitates robust implementation strategies informed by implementation science. By developing a comprehensive understanding of the barriers and facilitators, and harnessing frameworks like the CFIR and KTA Cycle, we can facilitate the systematic adoption of AI in healthcare. As AI continues to evolve, it is imperative that its deployment not only focuses on technical feasibility but also on practical usability, ensuring that AI solutions enhance patient outcomes and optimize healthcare delivery on a large scale. This requires a concerted effort from all stakeholders involved, creating an environment where innovation is not only pursued but effectively integrated into the fabric of healthcare delivery.

References

1. Lisacek-Kiosoglous AB, Powling AS, Fontalis A, Gabr A, Mazomenos E, Haddad FS. Artificial intelligence in orthopaedic surgery. Bone Joint Res 2023;12(7):447-54. doi: 10.1302/2046-3758.127.BJR-2023-0111.R1. PMID: 37423607; PMCID: PMC10329876.
2. Darzi A. Independent investigation of the National Health Service in England. September 2024. Available at: <https://assets.publishing.service.gov.uk/media/66f42ae630536cb92748271f/Lord-Darzi-Independent-Investigation-of-the-National-Health-Service-in-England-Updated-25-September.pdf>.
3. Balas EA, Boren SA. Managing clinical knowledge for health care improvement. Yearb Med Inform 2000;(1):65-70.
4. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019;17(1):195.
5. Sharma S. Artificial intelligence for fracture diagnosis in orthopedic X-rays: current developments and future potential. SICOT J 2023;9:21. doi: 10.1051/sicotj/2023018.
6. Breu R, Avelar C, Bertalan Z, Grillari J, Redl H, Ljuhar R, Quadlbauer S, Hausner T. Artificial intelligence in traumatology. Bone Joint Res 2024;13(10):588-95. doi: 10.1302/2046-3758.1310.BJR-2023-0275.R3. PMID: 39417424; PMCID: PMC11484119.

7. Liu X, Faes L, Calvert MJ, Denniston AK, Frost H. Artificial intelligence for the detection of fractures on plain radiographs: a systematic review and meta-analysis. *Eur Radiol* 2020;30(2):523–36.
8. Kuo RYL, Harrison C, Curran TA, Jones B, Freethy A, Cussons D, Furniss D. Artificial intelligence in fracture detection: a systematic review and meta-analysis. *Radiology* 2022;304(1):50–62.
9. Olczak J, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, Gordon M. Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop* 2017;88:581–86.
10. Urakawa T, Tanaka Y, Goto S, Matsuzawa H, Watanabe K, Endo N. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. *Skeletal Radiol* 2018;48:239–44.
11. Chung SW, Han SS, Lee JW, Oh KS, Kim NR, Yoon JP, Kim Y. Automated detection and classification of the proximal humerus fracture by using a deep learning algorithm. *Acta Orthop* 2018;89:468–73.
12. Langerhuizen DWG, Janssen SJ, Mallee WH, van den Bekerom MPJ, Ring D, Kerkhoffs GMMJ, Doornberg JN. What are the applications and limitations of artificial intelligence for fracture detection and classification in orthopaedic trauma imaging? A systematic review. *Clin Orthop Relat Res* 2019;477(11):2482–91.
13. Pauling C, Kanber B, Arthurs OJ, Shelmerdine SC. Commercially available artificial intelligence tools for fracture detection: the evidence. *BJR Open* 2023;6(1):tzad005. doi: 10.1093/bjro/tzad005. Erratum in: *BJR Open* 2024;6(1):tzae004.
14. Bergh C, Wennergren D, Möller M, Brisby H. Fracture incidence in adults in relation to age and gender: a study of 27,169 fractures in the Swedish Fracture Register in a well-defined catchment area. *PLoS One* 2020;15(12):e0244291.
15. Cooper C, Dennison EM, Leufkens HG, Bishop N, van Staa TP. Epidemiology of childhood fractures in Britain: a study using the general practice research database. *J Bone Miner Res* 2004;19(12):1976–81.
16. Cepela DJ, Tartaglione JP, Dooley TP, Patel PN. Classifications in brief: Salter-Harris classification of pediatric physeal fractures. *Clin Orthop Relat Res* 2016;474:2531–37.
17. Shrader MW, Campbell MD, Jacofsky DJ. Accuracy of emergency room physicians' interpretation of elbow fractures in children. *Orthopedics* 2008;31(12).
18. Li W, Stimec J, Camp M, Pusic M, Herman J, Boutis K. Pediatric musculoskeletal radiographs: anatomy and fractures prone to diagnostic error among emergency physicians. *J Emerg Med* 2022;62(4):524–33.

19. Pinto A, Berritto D, Russo A, Riccitello F, Caruso M, Belfiore MP, Grassi R. Traumatic fractures in adults: missed diagnosis on plain radiographs in the emergency department. *Acta Biomed* 2018;89(1-S):111–23.

20. Dunker D, Collin D, Göthlin JH, Geijer M. High clinical utility of computed tomography compared to radiography in elderly patients with occult hip fracture after low-energy trauma. *Emerg Radiol* 2012;19(2):135–39

21. Lex JR, Di Michele J, Koucheiki R, Pincus D, Whyne C, Ravi B. Artificial intelligence for hip fracture detection and outcome prediction: a systematic review and meta-analysis. *JAMA Netw Open* 2023;6(3):e233391.

22. Anderson GH, Jenkins PJ, McDonald DA, Van Der Meer R, Morton A, Nugent M, Rymaszewski LA. Cost comparison of orthopaedic fracture pathways using discrete event simulation in a Glasgow hospital. *BMJ Open* 2017;7(9):e014509.

23. Department of Health Audit Commission. Outpatients – review of national findings. Department of Health 2003.

24. Jenkins P, Gilmour A, Murray O, et al. The Glasgow Fracture Pathway: a virtual clinic. *BJJ News* 2014;2:22–24.

25. Guly HR. Diagnostic errors in an accident and emergency department. *Emerg Med J* 2001;18:263–69.

26. Pfeifer R, Pape HC. Missed injuries in trauma patients: a literature review. *Patient Saf Surg* 2008;2:20.

27. NHS Resolution. Annual report and accounts 2023/24. 2024. Available at: https://resolution.nhs.uk/wp-content/uploads/2024/07/NHS-Resolution-Annual-report-and-accounts_23-24_Access.pdf.

28. Voets MM, Veltman J, Slump CH, Siesling S, Koffijberg H. Systematic review of health economic evaluations focused on artificial intelligence in healthcare: the tortoise and the cheetah. *Value Health* 2022;25:340-9. doi: 10.1016/j.jval.2021.11.1362. PMID: 35227444.

29. Sharma S. Artificial intelligence for fracture diagnosis in orthopedic X-rays: current developments and future potential. *SICOT J* 2023;9:21.

30. Gao C, Hu C, Qian Q. Artificial intelligence model system for bone age assessment of preschool children. *Pediatr Res* 2024. <https://doi.org/10.1038/s41390-024-03282-5>

31. Thompson P, Khattak M, Joseph PJ, Perry DC, Cootes TF, Lindner C. Automating radiological measurements of the hip in children with cerebral palsy. *Bone Joint J* 2025;107-B(1):124-32. doi:10.1302/0301-620X.107B1.BJJ-2024-0894.

32. Shah AK, Lavu MS, Hecht CJ 2nd, Burkhart RJ, Kamath AF. Understanding the use of artificial intelligence for implant analysis in total joint arthroplasty: a systematic review. *Arthroplasty* 2023;5(1):54. doi: 10.1186/s42836-023-00209-z. PMID: 37919812; PMCID: PMC10623774.
33. Ormond MJ, Clement ND, Harder BG, Farrow L, Glester A. Acceptance and understanding of artificial intelligence in medical research among orthopaedic surgeons. *Bone Jt Open* 2023;4(9):696-703. doi: 10.1302/2633-1462.49.BJO-2023-0070.R1. PMID: 37694829; PMCID: PMC10494473.
34. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation frameworks for artificial intelligence translation into health care practice: scoping review. *J Med Internet Res* 2022;24(1):e32215.
35. Wang Y, Wong EL, Nilsen P, Chung VC, Tian Y, Yeoh EK. A scoping review of implementation science theories, models, and frameworks—an appraisal of purpose, characteristics, usability, applicability, and testability. *Implement Sci* 2023;18(1):43.
36. Nilsen P, Svedberg P, Neher M, Nair M, Larsson I, Petersson L, Nygren J. A framework to guide implementation of AI in health care: protocol for a co-creation research project. *JMIR Res Protoc* 2023;12(3):e50216.
37. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci* 2015;10:53.
38. May CR, Cummings A, Girling M, Bracher M, Mair FS, May CM, et al. Using normalization process theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review. *Implement Sci* 2018;13(1):80.
39. Farrow L, Meek D, Leontidis G, Campbell M, Harrison E, Anderson L. The Clinical Practice Integration of Artificial Intelligence (CPI-AI) framework. *Bone Joint Res* 2024;13(9):507-12. doi: 10.1302/2046-3758.139.BJR-2024-0135.R1. PMID: 39288942; PMCID: PMC11407877.
40. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, Robinson N. Lost in knowledge translation: time for a map? *J Contin Educ Health Prof* 2006;26(1):13-24.

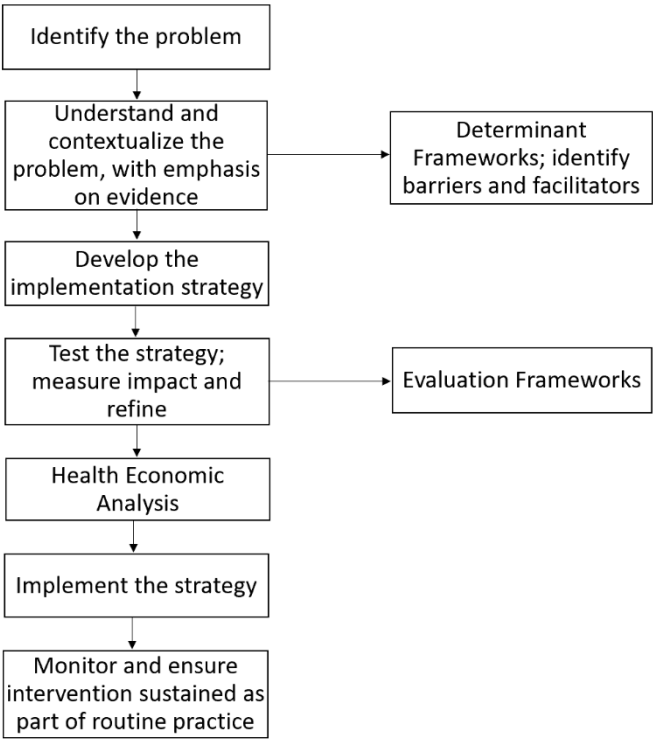


Figure 1 - Novel Technology Implementation Blueprint