The Heimdall framework for supporting characterisation of learning health systems

Scott McLachlan  
Queen Mary University of London, London, UK

Henry W. W. Potts  
University College London, London, UK

Kudakwashe Dube  
Massey University, Auckland, New Zealand

Derek Buchanan  
Massey University, Auckland, New Zealand

Stephen Lean  
Massey University, Auckland, New Zealand

Thomas Gallagher  
University of Montana, Missoula, MT

Owen Johnson  
University of Leeds, Leeds, UK

Bridget Daley  
Blizzard Institute, Queen Mary University of London, London, UK

William Marsh  
Queen Mary University of London, London, UK

Norman Fenton  
Queen Mary University of London, London, UK

ABSTRACT

Background  There are many proposed benefits of using learning health systems (LHS), including improved patient outcomes. There has been little adoption of LHS in practice due to challenges and barriers that limit adoption of new data-driven technologies in healthcare. We have identified a more fundamental explanation: the majority of developments in LHS are not identified as LHS. The absence of a unifying namespace and framework brings a lack of consistency in how LHS are identified and classified. As a result, the LHS 'community' is fragmented, with groups working on similar systems being not aware of each other’s work. This leads to duplication and the lack of a critical mass of researchers necessary to address barriers to adoption.

Objective  To find a way to support easy identification and classification of research works within the domain of LHS.

Method  A qualitative meta-narrative study focusing on works that self-identified as LHS was used for two purposes. First, to find existing standard definitions and frameworks using these to create a new unifying framework. Second, seeking whether it was possible to classify those LHS solutions within the new framework.

Results  The study found that with apparently limited awareness, all current LHS works fall within nine primary archetypes. These findings were used to develop
a unifying framework for LHS to classify works as LHS, and reduce diversity and fragmentation within the domain.

**Conclusion** Our finding brings clarification where there has been limited awareness for LHS among researchers. We believe our framework is simple and may help researchers to classify works in the LHS domain. This framework may enable realisation of the critical mass necessary to bring more substantial collaboration and funding to LHS. Ongoing research will seek to establish the framework’s effect on the LHS domain.

**Keywords:** electronic health records, learning health systems, learning healthcare systems, precision medicine

---

**INTRODUCTION**

Learning health systems (subsequently referred to as LHS) are defined by the Institute of Medicine (IoM) as systems in which alignment of scientific and cultural tools lead to knowledge generation to improve healthcare as a result of daily practice.\(^1\) Since LHS were conceptualised in 2007, they have been the focus of increasing research attention.\(^2\)–\(^6\) The opportunity and promise of LHS has resulted in texts presenting collections of LHS-specific research\(^1,7\) and creation of a new journal, LHS,\(^8\) and new courses of academic study.\(^9,10\)

We believe it could be the most significant development in healthcare since the advent of evidence based medicine (EBM) and electronic health records (EHRs) that support EBM.\(^14\)–\(^16\) EHR have existed for more than 40 years\(^11\)–\(^13\) and organisations that implemented EHR discovered reductions in costs, clinical testing and patterns of repeated and sometimes unnecessary prescriptions. Enhanced co-ordination and communication between clinicians was seen to improve the quality of patient care.\(^12,17\)–\(^20\)

Despite the benefits, early EHR systems were considered expensive, focused on information gathering rather than improving healthcare.\(^20\) Development lacked clinical input, existed as multiple stand-alone systems, experienced slow adoption, suffered from trust and data quality issues, claims that systems increase or exacerbate risk for errors, and concerns over patient privacy and security.\(^18\),\(^20\) All of these issues are still seen as unresolved barriers to adoption of EHR.\(^18,21\)–\(^27\) Despite this, EHRs are the foundation for LHS. Efforts towards LHS, coupled with proposed changes to legislation, policy and the ethics of how clinicians engage with clinical datasets suggest an entirely new dimension to EHR. One in which they are used collectively as ‘big data’ and focused using individual patient’s attributes to identify causes and optimal treatments strategies for disease.

Descriptions developed in IoM reports are the basis of most author’s definitions and descriptions of LHS.\(^28\)–\(^33\) Medical information systems that can be predictive, preventative, personalised and participatory represent the core principles of 4P medicine.\(^34\) According to the IoM, these systems have the potential to identify groups at greatest risk of complications for purposes of targeting interventions.\(^7\) In parallel, maturing technologies such as large datasets, machine learning, and enhanced processing power further enable the concept of LHS.\(^2,34,35\)

The IoM promoted EBM as the primary driver for LHS,\(^7\) yet their definition fails to describe LHS attributes which contribute to quality, safety, efficiency and effectiveness of patient care.\(^36\) The four fundamental attributes\(^37\) listed in Figure 1 provide tangible metrics to compare and contrast LHS efforts. These attributes were not included in the IoM definition, but are widely found in LHS. They clarify the use of large EHR datasets as the source of knowledge for LHS achieving the goal of improving quality in individual patient care.

There are numerous examples of proposed benefits of LHS. For clinicians, these include assessing which laboratory or imaging tests may be more diagnostic given a patient’s presenting symptomology\(^38\); reducing risk from prescribing

---

1. An organizational architecture that facilitates formation of communities of patients, families, front-line clinicians, researchers, and health system leaders who collaborate to produce and use big data;
2. Large electronic health and health care data sets (big data);
3. Quality improvement for each patient at the point of care brought about by the integration of relevant new knowledge generated through research; and
4. Observational research and clinical trials done in routine clinical care settings.

Adapted from: [37]

**Figure 1 Four elements of an LHS**
errors\textsuperscript{31} and increasing awareness of pharmacogenetics.\textsuperscript{39} It is claimed that patients would benefit from advanced knowledge developed from the experiences and diagnoses of past patients, which saves time and reduces costs.\textsuperscript{38,40,41} A learning healthcare organisation culture supports EBM, while successful integration of research into practice is what enables it.\textsuperscript{42} The ability to use technology to record, compare, contrast and present information in almost real-time enhances the input, analysis and decision phases of the learning lifecycle. Alternatively, it is said that the financial burden to implement and support health technology\textsuperscript{43–45} along with a persistent need for data and systems standardisation,\textsuperscript{46,49,50} interoperability,\textsuperscript{49,51} and integration\textsuperscript{49,51} have all acted as barriers to broad LHS adoption.

In our group’s recent letter to the editor of this journal,\textsuperscript{52} we demonstrated the lack of awareness and barriers for researchers to appropriately identify their efforts as LHS solutions. We believe that this results from a number of significant problems in the domain. The lack of adequate classification and standardisation results in groups working on comparable systems not identifying their efforts as LHS, and may be the cause of unnecessary duplication of efforts and the observable lack of collaboration. The absence of a unifying framework means the domain is yet to generate a necessary critical mass, limiting efforts to resolve barriers and challenges to the adoption of LHS and constraining funding availability. We were only able to identify one primary article that attempted to consolidate and analyse the current state of knowledge in LHS.\textsuperscript{36} We extend that effort drawing on a larger collection of works to establish a unifying framework for LHS.

**METHOD**

Our study sought first to define a comprehensive framework and taxonomy for LHS, and then to demonstrate its application to self-identified literature from the LHS domain. The literature search used identical plain language search terms as Foley and Vale\textsuperscript{36}: ‘LHS’ and ‘learning healthcare system’ drawing articles from Scopia, Science Direct, PubMed, EBSCOhost, DOAJ and Elsevier (\(n = 1083\)). These works were all authored in the decade since the IoM’s initial LHS report.\textsuperscript{7} Figure 2 shows this literature, by year, in orange (the drop in 2017 is due to reporting only up to July); contrasted with literature identified by Foley and Vale\textsuperscript{36} shown in grey. The leap in publications in 2011 followed the first\textsuperscript{53} and second\textsuperscript{54} meetings of the Committee on the Learning Healthcare System in America, both proceedings published during 2011. More than 50\% of LHS publications were generated since 2013.

We undertook this review following the identification of seminal sources approach of meta-narrative reviewing.\textsuperscript{27,55} An initial read of abstracts was used to reject duplicates and papers not related to the central topic, including those using the search term in context of learning in the academic or education sense.\textsuperscript{56,57} This reduced the collection source pool (\(n = 542\)). Conclusions and methods were reviewed, seeking to reject papers that did not present or propose an LHS; for example, those exploring the medicolegal, ethical or societal aspects of LHS.\textsuperscript{31,58,59} The resulting core pool (\(n = 230\)) was then comprehensively reviewed. Of these, 53\% proposed a potential solution compared to 47\% that presented an existing solution.

We used content and thematic analysis\textsuperscript{60} to recognise and classify LHS uses while also identifying common barriers and thematic concepts for investigation. Formal concept analysis\textsuperscript{61} was used to identify the frequency and interrelationships between the identified concepts. The elements of both analysis methods were identified inductively during the first full reading of the core pool of literature and used to develop spreadsheets for analysis. A second full reading was performed to data mine the literature and populate spreadsheets for further analysis during framework development. Table 1

![Figure 2 LHS publications by year](https://example.com/figure2.png)
**RESULTS**

**Taxonomy for LHS**

We found only three papers\textsuperscript{36,62,63} that proposed classification systems for LHS. Surveillance and Comparative Effectiveness Research were the only types common to all three. Figure 3 unifies the knowledge identified from all three papers into a taxonomy of nine LHS classification types, indicating abbreviations (initials) and the primary reference for each.

Cohort identification (CI) seeks patients with similar attributes, used to determine the feasibility of studies and quantify numbers of potential patients that may be helped.\textsuperscript{62} CI is also the first operational step of most other LHS types.

Positive deviance (PD) uses outcome data to benchmark clinical care. PD identifies elements of safer, more effective, timely and patient-centred care, recognising beneficial behaviours for incorporation into another clinician’s practice.\textsuperscript{64} PD can also identify common traits of patients benefiting from a treatment, using these to identify others who may benefit from the same intervention.\textsuperscript{62}

Negative deviance (ND) identifies instances of sub-optimal care outcomes. ND presupposes some particular clinical behaviour negatively impacted patient care and the resulting outcome. The clinician critically evaluates the care provided, investigating causes for sub-optimal results.\textsuperscript{63}

Predictive patient risk modelling (PPRM) uses patterns discovered in patient datasets to identify cohorts at higher risk for future adverse events. PPRM can use routine health data to identify ‘triple fail’ events; where treatment fails to improve patient care experiences, population health, or lower health-care costs.\textsuperscript{65}

Predictive care risk and outcome model (PCROM) algorithms identify situations of high risk for ‘unsafe’, ‘delayed’ or ‘inefficient’ care, providing estimates of the effectiveness of different interventions.\textsuperscript{36} Geisinger Health Systems have incorporated PCROM approaches into clinical software, identifying when spikes in hospital activity or patient non-attendance may occur.\textsuperscript{66}

Clinical decision support systems (CDSS) are active knowledge systems where two or more characteristics of the patient are matched to computerised knowledge bases with algorithms generating patient-specific treatment recommendations.\textsuperscript{67–69}

Comparative effectiveness research (CER) compares interventions and outcomes within an EHR dataset to determine the most effective treatment, using a method considered more efficient than randomised control trials.\textsuperscript{36} CER isolates patients with similar attributes to the current patient,
returning knowledge on treatments that deliver optimum health outcomes.70

Intelligent assistance (IA) uses data sources to automate routine processes such as prepopulating pathology orders and clinical notes, or summarising patient case notes prior to consultations.36

Surveillance (S) monitors EHR data for outbreaks of disease (e.g., measles) or treatment issues (e.g., contaminated medicines or increased frequency for post-surgical infections). Examples observed include health and demographic surveillance systems used in sub-Saharan Africa.71

The Heimdall-integrated LHS framework

Just as the Norse God Heimdall was said to be the son of nine mothers, we started from our nine LHS classifications to develop the integrated LHS framework in Figure 4. The diagram’s conical structure demonstrates the use of technology (large datasets and processing systems) to record, store, index and present information that flows into and improves the learning processes used in EBM, focusing clinical practice towards delivery of precision medicine (PM). This enables the learning healthcare organisation to engage in decisions individualised to match unique patient characteristics.

PM results from enhancing the generalised population health approach using attributes in the EHR to constrain analysis for diagnosis and treatment options to cohort predominately matching the presenting patient’s profile. As the clinician enters attributes about the current patient, the speed and accuracy of decisions increase as illustrated by arrows in Figure 4. LHS draw knowledge from a reducing cohort whose attributes predominantly match the current patient as illustrated by the circular design. In examining a cross-section of Figure 4, the larger white circle represents the entire population used to select the most effective common treatment. The light grey circle reduces that population to those who share some basic attributes with the patient that would normally be identified in the slower learning organisation approach of EBM. The inner dark circle further reduces the population to a significant cohort with clinical, genetic and socioeconomic attributes predominately matching the patient at the centre.

The interrelationship between the Heimdall framework and our taxonomy is shown down the right side of the conical portion of the diagram. While LHS are technology solutions, the majority operate in the context of either the treatment provider’s learning organisation, or the clinician’s primary patient-facing role.

Within this framework, we incorporate the concept of a clinical lifecycle as shown in Figure 5 and adapted from multiple works in this review.63,66,70 The right side of the diagram represents largely clinician-driven aspects, while the left side identifies those aspects where LHS technology delivers improvements. The more challenging barriers to LHS regarding data quality, interoperability and standardisation all result from activities on the right side of the lifecycle.

This cycle is repeated, both for surveilling the proposed transformation and to seek further items of knowledge.63,66,70 LHS engender a close relationship between care, research and knowledge translation, aimed at providing a platform for integrating various data to better understand patients.72 LHS are commonly described using an iterative lifecycle similar to many other EBM processes where: (a) patient data is collected by clinicians; (b) aligned, transformed and amalgamated into larger data sets; (c) a problem is defined and (d) analysis performed; (e) with evidence data returned and (f) made into new knowledge, that is, (g) used to transform clinical practice.

The taxonomy and Heimdall framework provided a toolkit for characterising LHS literature in terms of the following thematic and analytical aspects to assess whether the LHS demonstrates:

- Taxonomic consistency – conforms to the taxonomy.
- Patient focused – ensures personalised health care, known as PM.
- Technology usage – uses health IT with big data, machine learning algorithms and automation.
- Decision support – near real-time support to clinicians, bringing recent scientific advances, machine learning and EBM together at point-of-care.
- Application of LHS – goes beyond selecting ‘most likely treatment for a population’, to selecting the ‘most applicable treatment for an individual’.
- Barriers and further observations – challenges limiting implementation.

What follows is a survey and discussion using the Heimdall framework to identify and resolve key questions from within
the domain of LHS. Answers to these questions were resolved through applying the framework to the broader literature cohort.

**Validation of the LHS taxonomy**

To validate our taxonomy, proposed or presented solutions were reviewed and classified using the taxonomic descriptions. Some proposed solutions were incompletely described, but we found that the intention of the authors was always clear from the information presented. All LHS solutions conformed easily to one of our identified taxonomic types. We believe this validates the taxonomy as we have presented. Our validation of the taxonomy also found that CER were the most prevalent type of LHS presented in the literature, as identified in Table 2. CER was followed at some distance by stand-alone CI, although it should be noted that this classification type had not been identified by either Deeny and Steventon or Foley and Vale.

**DISCUSSION**

A number of themes emerged during this analysis and are discussed in the following section, which includes topical analysis to investigate their effect on the LHS domain.

**Patient focused**

While clinicians argue they always practiced patient-focused medicine, normal clinical practice follows population medicine-based EBM. PM extends diagnostic practices with profiling techniques and therapies tailored to the individual. Patient focus is a key dimension that LHS improve. PM approaches can be retrospective, as in CER, and prospective, when genotyping for treatment selection. Patient-centred care encourages data use in optimising care for individuals. Aggregated patient records enable LHS to identify cohorts similar to the patient. LHS is an efficient tool for integrating PM into practice. As the clinician enters attributes about the patient, the LHS refines a cohort of prior similar patients. It assesses the treatments they received to recommend one most likely to produce an optimal outcome.

**Technology usage**

The focus for health IT has shifted from issues of adoption to identifying how to best use technology to improve healthcare delivery and outcomes. This shift is significant in creating LHS and elevates issues in EMR/EHR interoperability, data standardisation and quality that must be resolved if LHS are to be truly practical and ubiquitous. Health IT’s ability to improve healthcare service delivery quality and efficiency is recognised. Enabling necessary data flow and integration of data sources are key abilities technology can deliver, representing core requirements to enabling LHS. Integration of learning into technology is observed in every LHS solution.
Decision support
Healthcare providers evolved from considering health IT as a billing and documentation facilitator, to contemplating its active participation and capabilities to answer complex questions in care delivery.85 LHS bring opportunities for improving speed and efficiency of clinical decision support.83,85,86 LHS solutions are context-sensitive, incorporating CI and risk modelling in real time to identify interventions for improving individual patient outcomes.85,88 LHS have potential to rapidly perform retrospective comparative effectiveness trials, evaluating treatment options against each other where they have been provided to similar patients.88,89 In contrast with randomised clinical trials, LHS is considered safer, and engenders greater confidence in accuracy of the treatment choice.88,89

Application of LHS
Clinical epidemiology is an example of learning healthcare. EBM evolved from clinical epidemiology: statistically identifying the optimal treatment which becomes best practice for that condition.90,91 Conversely, the focus for PM is selecting from available interventions the treatment that will best serve the individual. The primary driver towards population medicine was economic: maximising benefit while minimising cost, harm and waste.90 While meant for benefiting the individual patient, population medicine has disadvantages in that the individual’s best interests may conflict with those of the population and it is difficult to reconcile the two.76,90 Individual patients may be denied higher priced precision interventions in favour of lower cost population-optimised interventions.90 The Heimdall framework demonstrates that LHS focus healthcare using population medicine and EBM directly onto the presenting patient. While EBM selects one treatment for all patients, LHS produces a cohort with attributes similar to the presenting patient, identifying the treatment most likely to be effective for this individual patient. LHS in this way delivers PM.

Barriers and further observations
Most authors discuss barriers to implementation. The most common are cost,32,92,93 data interoperability and standardisation,94–96 poor data quality and integrity,93,97,98 informed consent and ethics review complications,99–101 privacy and security issues76,95 and slow technology adoption.95,102,103 These issues are seen in the same context for adopting EHR/EMR. This suggests LHS are inheriting problems from the EHR/EMR on which they depend.

CONCLUSION
LHS represent a significant improvement on the present learning organisation, evidence-based practice and population medicine approaches. LHS improve the focus on the individual patient, bringing efficient and expedient PM solutions. LHS are a significant evolution to EBM, and the natural next step in realising the benefits that were expected from implementing EHR/EMR. However, the lack of taxonomy for classifying and describing LHS may be a significant reason for fragmented and duplicated research and solutions in LHS that has impeded widespread adoption. Many authors presenting solutions fail to identify them as LHS. Our research has presented a taxonomy and framework to address this problem, and may help address the challenges in realising all that LHS promise.

Funding
Scott McLachlan, William Marsh and Norman Fenton acknowledge support from the EPSRC under project EP/P009964/1: PAMBAYESIAN: Patient Managed decision-support using Bayes Networks. Henry W. W. Potts has received consultancy fees from Crystallise and The HELP Trust, and funding from myownteam and Shift.ms.

Competing Interests
None of the authors identified a competing interest relevant to this research.

Contributors
William Marsh raised the topic, leading to Scott McLachlan performing the primary research and preparing the first draft of this paper. Kudakwashe Dube proposed formation of the framework, Owen Johnson developed the initial framework which was further enhanced by Derek Buchanan from the clinician’s perspective. Derek Buchanan and Bridget Daley provided review and input based on clinical experience. The initial method was proposed by Scott McLachlan, extended and significantly improved by Henry W. W. Potts, with Stephen Lean, Kudakwashe Dube and Thomas Gallagher providing input to the survey structure and framework. Thomas Gallagher, Kudakwashe Dube, Norman Fenton, Bridget Daley, Stephen Lean and Henry W. W. Potts provided editorial review. Norman Fenton and William Marsh supervised the research. All authors commented on and approved the final draft.

REFERENCES
improvement tapestry: learning health system brings together health IT data stakeholders to share knowledge and improve healthcare.


Cucin V. Embedding data provenance into the learning health system to facilitate reproducible research. Learning Health Systems 2017;1(2).


92. Boonstra A and Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Services Research 2010;10(1):231.


