electronics

MDPI

Article

A Solution for the Health Data Sharing Dilemma: Data-Less and Identity-Less Model Sharing Through Federated Learning and Digital Twin-Assisted Clinical Decision Making †

Nilmini Wickramasinghe * and Nalika Ulapane

School of Computing, Engineering and Mathematical Sciences, La Trobe University, Melbourne, VIC 3086, Australia; n.ulapane@latrobe.edu.au or nalika.ulapane@gmail.com

* Correspondence: n.wickramasinghe@latrobe.edu.au or nilmini.work@gmail.com

† This paper is a revised and expanded version of a paper entitled "Digital Twins of Patients: A Liquid Neural Network Interpretation", which was presented at ACIS 2024, Canberra, Australia, 4–6 December 2024.

Abstract: Digital twins are essentially digital replicas of physical entities. Their usage is becoming more common across various industries, including healthcare. However, the implementation of digital twins in healthcare is uniquely challenging. This is partly because of the sensitive nature of health data and privacy concerns. These concerns limit health data accessibility and shareability. This paper attempts to address this challenge of health data sharing. We propose a novel approach that leverages federated learning, model sharing, and digital twin-assisted clinical decision making. Our approach ensures that health data are kept federated with healthcare providers. Healthcare providers train machine learning models on their own data. Then, instead of sharing the data, the trained models are shared. This is enabled via an arrangement like a private blockchain that is accessible to subscribed healthcare providers. This approach allows healthcare providers to access and use machine learning models for clinical decision support without compromising sensitive data about patients. Certain information about machine learning models will be shared. These include indicators such as the sample size on which a model has been trained on, validation metrics, and model accuracy. Such information assists other healthcare providers in selecting the most effective models. We demonstrate the efficacy of this approach through a case study on chronic disease management (e.g., cancer) using Liquid Neural Networks. Our results show how federated learning and model sharing can enhance clinical decision making and improve patient outcomes while ensuring the privacy of data.

Keywords: blockchain; clinical decision support; digital twin; federated learning; liquid neural network; machine learning

check for updates

Academic Editors: Stanley C. Ahalt and George A. Tsihrintzis

Received: 21 December 2024
Revised: 20 January 2025
Accepted: 6 February 2025
Published: 10 February 2025

Citation: Wickramasinghe, N.; Ulapane, N. A Solution for the Health Data Sharing Dilemma: Data-Less and Identity-Less Model Sharing Through Federated Learning and Digital Twin-Assisted Clinical Decision Making. *Electronics* **2025**, *14*, 682. https://doi.org/10.3390/electronics14040682

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Digital twins, i.e., the digital replicas of physical entities, have been emerging in several industries [1,2]. Advancements on the internet of things have enabled them to come to practice coupled with Industry 4.0. Along with this trend, some early works have also emerged, attempting to adapt the concept of digital twins to the healthcare sector as well, with particular focuses on creating digital twins of humans (involving genomics) [3], or certain aspects of humans like chronic disease progression [4,5], and also healthcare processes and workflows [6,7]. However, making use of digital twins in practice requires access to data, ideally in real time. Making data accessible in the healthcare sector though, especially when it involves the sensitive data of patients, becomes

problematic. The sensitive nature of the data of patients alongside the privacy concerns inhibits the accessibility and the shareability of healthcare data. This in turn inhibits the implementation of digital twins of patients in healthcare. Through this paper we attempt to present a solution to address the underlying dilemma of health data sharing. We attempt to do so by answering the following research question: How might the barrier of sharing health and medical data of patients be overcome through federated learning, model sharing, and digital twin-assisted clinical decision making?

To answer the aforesaid research question, we propose a novel approach that ensures the health data remain with the healthcare providers, for instance, with the hospitals. This means the health data remain federated. Each healthcare provider employs machine learning to train models on their own data. Then, instead of sharing the data, the models are shared, for instance, via something like a private blockchain dedicated for healthcare providers [8–10]. This private blockchain will be accessible to healthcare providers who have subscribed to it. This approach allows healthcare providers to access machine learning models produced by different healthcare providers.

The healthcare providers will have access to certain information about the available machine learning models, for example, indicators like the sample size on which a model has been trained on, the sample size on which a model has been validated on, the recency of a model, and indicators regarding the accuracy of a model. Such information will help a skilled data scientist or a statistician within the healthcare provider's workforce to select the most powerful and up-to-date machine learning models that are available. These models can then be deployed in a backend to facilitate a frontend for clinicians. The clinician-facing frontend will typically function as an intuitive and easy-to-use clinical decision support tool—something modern clinicians are familiar with. As such, through this paradigm, healthcare providers will have access to machine learning models that are usable as clinical decision support tools. But the data on which those models have been trained will not be accessible to external parties. Thus, the healthcare data remains federated. The healthcare data do not have to be shared. The patient data and identity will therefore remain secure with the healthcare providers from whom a patient seeks care from. When it comes to creating and using digital twins of patients, some of the vital and sensitive data elements of which the privacy and anonymity must be preserved include genomic data, biometrics, pathological data, physiological data, medical scans, patient's medical history, treatment history, patient reported outcomes, and certain patient behavioral and visual elements that might be captured through any wearable or environmental sensors [11].

If desired, healthcare providers who access machine learning models produced by other providers can validate these models against their own data and select the models with the highest validation scores. Such a model selection approach ensures some action of intelligent model selection for clinical decision support. The proposed method thus enhances the privacy of data. It also reduces the risk of data breaches. Moreover, collaborative learning and decision making are promoted without compromising sensitive data.

As a case study example, we demonstrate the efficacy of this approach through models trained for chronic disease management, for example, cancer. It is well known that accessing real-world healthcare data is challenging. Even if they are accessible, they would often be incomplete and insufficient. Therefore, we carry out this work using simulated data. As the machine learning technique, Liquid Neural Networks (LNNs) are used because we attempt to forecast chronic disease progression, and LNNs have emerged recently as an adaptable and robust means for handling dynamic data [12–16], making them suitable for this application.

This paper makes a twofold contribution: a contribution to theory and a contribution to practice. This paper contributes to theory by demonstrating how LNNs can be applied in

healthcare settings targeted at chronic disease management. As a case study example, modeling cancer tumor radius is attempted. This paper contributes to practice by proposing an approach to tailor and use a blockchain-like architecture to share machine learning models along with vital information among healthcare providers, instead of sharing sensitive healthcare data. Healthcare providers can learn machine learning models from federated data that they possess and then can securely share models, ensuring transparency and traceability, while raw health data are kept private.

Our results demonstrate the power of sharing machine learning models in healthcare, even if the data are kept federated. Our simulated results show how significantly model sharing can assist clinical decision making and improve patient outcomes. The approach we propose thus enables a secure and efficient means to leverage the capabilities of advanced machine learning, model sharing and collective intelligence in decision making in healthcare, while data are kept private and federated. This paper thus presents one of the first works of its kind, showing the potential of federated learning and model sharing through means like private blockchains in the healthcare sector [8].

The remainder of this paper goes as follows: Section 2 presents a brief review of the related literature; Section 3 presents the methodology and results that demonstrate the effect of model sharing in healthcare through simulated data; Section 4 discusses a blockchain-like architecture to enable model sharing in healthcare practice; in Section 5, a discussion is presented about the implication; and, in Section 6, the paper is concluded, summarizing the implications and directions for future work.

2. Literature Review

2.1. Clinical Decision Making and Clinical Decision Support Systems (CDSSs)

Healthcare provision has evolved over time. This evolution has been influenced by advancements in medical knowledge as well as technology. As a result, over time, patient outcomes and the general efficacy of healthcare provision have improved over the past several decades. A core component that underpins the improvement of healthcare provision is accurate clinical decision making [17]. Clinical decision making too has evolved over time. In earlier times, clinical decision making was dominantly reliant on medical knowledge, experience, and competence of the healthcare professionals such as doctors and nurses. But, in more recent times, thanks to advancements in technology, it has been possible to incorporate data and information-based clinical decision support systems (CDSSs). These systems make use of healthcare-related data that are available along with data analytic techniques to provide evidence-based recommendations [18]. The introduction of CDSSs has thus enhanced the accuracy and effectiveness of clinical decision making.

2.2. Challenges in Sharing Data in Healthcare

Health data sharing and accessibility to health data are essential to enable best CDSSs. The access to rich healthcare datasets enables the development of machine learning models that can make evidence-based or data-driven predictions that can help clinical decision making. These predictions might encompass the likes of predicting the likely patient outcomes before a patient undergoes a certain intervention, predicting potential health risks to a person based on different aspects about them and recommending personalized treatment plans [19]. However, sharing sensitive health data to enable such predictions comes with challenges. These challenges include the likes of privacy concerns, regulatory constraints, and the risk of data breaches [20]. Such challenges pose barriers to sharing health data. This in turn limits the ability to develop effective and data-rich machine learning models and the ability to put these models into practice as CDSSs.

2.3. The Use of Federated Learning (FL) in Healthcare

One way to address the challenges associated with sharing healthcare data is by keeping health data federated. This means the data remain within the custodianship of individual healthcare providers. This approach reduces the risks associated with data sharing. Federated data allows federated learning (FL) [21]. FL means machine learning models that are trained across multiple decentralized datasets, or federated datasets. This allows developing machine learning models locally without compromising the privacy of data. FL has thus emerged as a promising approach to address the challenges associated with data sharing in healthcare.

A review in [22] has highlighted the potential of FL in healthcare. The ability to enhance data-driven insights while maintaining data security has been emphasized. This review has discussed various aspects of FL, including algorithms, the preservation of privacy, and several applications that demonstrate the potential of FL in different healthcare contexts.

Another systematic review [23] has explored the implementation of FL in healthcare. The authors have identified key challenges such as the heterogeneity of data, challenges associated with communication and connectivity, challenges associated with the convergence of machine learning models, and more. They have also highlighted several healthcare case studies where the use of FL has been explored. These case studies include applications in autism, cancer, COVID-19 detection, human activity recognition, emotion recognition, hospitals, patient mortality prediction, and sepsis.

2.4. The Use of Blockchain Technology in Healthcare

The blockchain technology [24] has also become an enabler for sharing information. It provides a secure and transparent platform for exchanging information. In healthcare, blockchain can be used to share machine learning models. This ensures that sensitive health data remain private, but required information including machine learning models can be shared through a trusted network where healthcare providers can subscribe, access, and validate where needed.

The review in [25] has studied blockchain applications in healthcare. The potential of blockchain technology to enhance the management of health data by improving security has been highlighted. The review has discussed various use cases. These include data management, supply chain management, and the management of clinical trials.

The systematic review in [26] has reported the ongoing research regarding blockchain in healthcare. The authors have identified several barriers to implementation. These barriers include the lack of interoperability among different databases and information systems in healthcare, the lack of standardization, the lack of scalability, the high volume of data, a lack of patient engagement and willingness to manage their own data, issues with data security and privacy, and a lack of incentives for the uptake of blockchain technologies. The authors have also presented several case studies that demonstrate the potential for integrating blockchain in healthcare. These case studies include the management of electronic medical records, healthcare-related supply chain management, biomedical research, education, elimination of the falsification of data and the under-reporting or exclusion of undesirable results in clinical research, remote patient monitoring, health insurance claims, and health data analytics.

2.5. Federated Learning with Blockchain to Address the Issues in Healthcare Data Sharing

Research discussed in the above subsections hints at the ability to integrate FL with blockchain technology. FL combined with blockchain offers a promising solution that allows healthcare providers to share machine learning models rather than sensitive health

data. This model-sharing approach enhances the privacy of data while enabling some degree of collaboration. Through this, there is potential for healthcare providers to create a secure and efficient framework for collaborative clinical decision making whilst addressing the critical issues of data privacy, security, and interoperability.

2.6. The Use of Liquid Neural Networks (LNNs)

We present a use case later in this paper, on how LNNs can be used to build data-driven models to assist clinical decision making. This subsection is therefore dedicated to briefly reviewing LNNs. Research regarding LNNs is a rapidly developing area in machine learning, and we draw insights from several recent references [12,13,15,16]. These studies show consensus that LNNs can adapt to temporal data. This adaptability stems from the design of LNNs, which is typically a system of linear first-order dynamical systems that are modulated by nonlinear interlinked gates [15]. This configuration allows LNNs to maintain stable and bounded behavior. Moreover, this configuration can be expressed as an arrangement of ordinary differential equations within a neural network architecture, thereby enhancing the performance in time-series prediction tasks [15].

A notable strength in LNNs has been identified as their ability to self-adapt to sudden changes in data patterns without having to be retrained [12]. This makes LNNs particularly valuable for applications that involve dynamic temporal data. Additionally, LNNs have been found to be able to learn and predict gradients or the rate of change in dynamic temporal data. This attribute has been beneficial for time-series predictions [13,15] including tasks such as traffic forecasting [12]. LNNs have also been used in various fields such as robotics, causality analysis [13], and even modeling liquid splash dynamics [16]. They are also being optimized to be useable in embedded systems that have limited computational resources [13].

Inspired by the ability of LNNs to predict gradients and their ability to adapt, later in this paper, we present a case example that we carried out to explore the potential of LNNs as digital twins of patients in healthcare, specifically, attempting to forecast patient progression in chronic diseases like cancer. We thereby demonstrate the feasibility of using LNNs as clinical decision support tools. They perform as sharable machine learning models that are learned through federated learning.

3. Methods

In this section, we present a simulated scenario through synthetic data that depicts the power of the approach that we propose to share machine learning models learned through federated data, while the data itself is kept private and unshared. A simulated scenario is performed because of the difficulty and, often, the insufficiency of real-world data that can be sourced for our analysis. We demonstrate how impactful it is when shared models are used for clinical decision making. For this, we consider a scenario involving two simulated hospitals. Suppose we name them Hospital A and Hospital B. Hospital A records the data of a cohort of patients treated by Hospital A, and, with that data, Hospital A learns machine learning models. These models are then made available to Hospital B. Suppose Hospital B accesses these models and uses them for clinical decision support. Thus, Hospital B will have access to clinical decision support tools that were produced by another healthcare provider, without having access to the data of other providers, thereby realizing the federated learning and model sharing approach we propose. Then, we demonstrate how the patient outcomes of Hospital B are impacted by this.

3.1. Clinical Scenario in Hospital A

Following our previous work [27], we start by supposing Hospital A receives an N, $N \in \mathbb{Z}^+$ number of patients ($N = 385$ is considered for modeling purposes following standard sample calculation principles for a size of unknown population size [28]). These patients are diagnosed with a certain type of cancer. Each of these patients get treated for their cancer at Hospital A by going through a maximum of n, $n \in \mathbb{Z}^+$ number of clinical encounters. For modeling purposes, we set $n = 6$ as a manageable and a realistic scenario of the number of treatment cycles for cancer treatment [29].

For a patient, the clinical encounters will be equally spaced in time. At the m^{th} clinical encounter ($m \in \mathbb{Z}^+, m \leq n$), the radius of a cancer tumor of the k^{th} patient ($k \in \mathbb{Z}^+$, $k \leq N$) is denoted by $y_m^{(k)}$, $y_m^{(k)} \in \mathbb{R}^+$, $\forall m, k$. Through our modeling, we attempt to model these cancer tumor radii.

The k^{th} patient receives a clinical intervention (such as a dosage of chemotherapy) at their m^{th} clinical encounter, and this clinical intervention is quantitatively denoted as $x_m^{(k)}$, $x_m^{(k)} \in \mathbb{R}^+$, $0.1 \leq x_m^{(k)} \leq 0.9$, $\forall m, k$. According to this scale, $x_m^{(k)} = 0.1$ is considered the minimum prescribable dosage and $x_m^{(k)} = 0.9$ is considered the maximum prescribable dosage $\forall m, k$.

It is assumed that the initial conditions for tumor radii are $0.5 \leq y_1^{(k)} \leq 0.9$, $\forall k$. It is also assumed that the initial treatment provided is $x_1^{(k)} = 0.5$, $\forall k$, assuming that the scaled value of 0.5 is a moderate intensity of treatment.

At this stage, suppose that Hospital A has minimal facility for data-based modeling. Therefore, at the subsequent clinical encounters, the intervention $x_m^{(k)}$, $m \geq 2$ is decided by the clinician to be a random number such that $0.1 \leq x_m^{(k)} \leq 0.9$, $\forall m, k$.

For the purpose of modeling, it is assumed that, under the above conditions, the tumor radii of all patients vary according to Equation (1), where $\epsilon_m^{(k)}$, $\epsilon_m^{(k)} \in \mathbb{R}$ is additive and random noise, such that $\left|\epsilon_m^{(k)}\right| \leq 0.05 y_m^{(k)}$, $\forall m, k$. The parameters $\alpha^{(k)}$, $\beta^{(k)}$ are such that $\alpha^{(k)} = 1$, $\forall k$, and $\beta^{(k)} \in \mathbb{R}^+$, $0.1 \leq \beta^{(k)} \leq 0.9$, $\beta^{(k)} \sim \mathcal{N}\left(0.5, 0.1^2\right), \forall k$. Thus, the value of $\beta^{(k)}$ is what dictates the uniqueness of a patient in this modeling example.

The noise $\epsilon_m^{(k)}$ is not measurable by the clinicians. The parameters $\alpha^{(k)}$, $\beta^{(k)}$ are also not known to the clinicians and are not measurable. A clinician only measures $y_m^{(k)}$ and prescribes $x_m^{(k)}$ for a given patient at any clinical encounter.

$$y_{m+1}^{(k)} = \alpha^{(k)} y_m^{(k)} + \beta^{(k)} \left(x_m^{(k)} - 0.5\right) + \epsilon_m^{(k)} \tag{1}$$

Suppose each patient is treated across $m = 1, 2, 3, \ldots, 5$ clinical encounters, the maximum being 5, i.e., n = 5 according to the aforesaid notation of n. At the final clinical encounter, i.e., at $n = 6$, the final patient outcome, i.e., the tumor radius $y_6^{(k)}$ is measured $\forall k$. Also suppose that the data of all patients are recorded in Hospital A and are available for subsequent modeling purposes.

3.2. Machine Learning Performed by Hospital A

Now suppose Hospital A performs some machine learning, using the data of these cancer patients with the intention of using the learned models as clinical decision support tools. For this, following standard sample size calculation principles assuming an unknown population size [28], Hospital A chooses the data of $N = 385$ patients to learn machine learning models.

The tumor radii are modeled as a time series. Given the typical dynamic and complex nature of the variation in tumor radii subject to different treatments across different patients, we decided to model the variation in tumor radii using neural networks. An adapting

neural network architecture was chosen with the aim of handling the dynamic nature of tumor data.

The chosen neural network architecture carries one hidden layer with r, $r \in \mathbb{Z}^+$ number of neurons. This is preceded by one input layer. A network is designed to take $2m$ inputs at the m^{th} clinical encounter of any patient. The network then models the tumor progression (i.e., tumor radius) as shown in Equation (3), using the commonly used hyperbolic tangent activation function $\tanh(t)$ for $t \in \mathbb{R}$ given in Equation (2) [30]. When modeling, we set r = 10, i.e., we set the hidden layer to have 10 neurons.

As such, this neural network adapts according to the clinical encounter m and outputs the rate of growth of the tumor for patient k expected until the subsequent clinical encounter (i.e., $m + 1$). The proposed neural network architecture thereby represents the attributes of adaptability [12,13] of an LNN, in terms of adapting according to the patient and the clinical encounter they are facing, and predicting the rate of change in a cancer tumor.

As shown in Equation (3), $u_i^{(m)}, v_i^{(m)}, w_j^{(m)}, b_1^{(m)}, b_2^{(m)} \in \mathbb{R}$ are model parameters. These parameters are learned from the data of the $N = 385$ patients described earlier. In our case (acting as Hospital A) we did the learning by splitting the data into training (50% of the data), testing (25% of the data), and validation (25% of the data) sets, to ensure the learned models are reliable while avoiding the chances of overfitting. Through this way of modeling, neural networks could be learned for each clinical encounter.

The approach followed by Hospital A is summarily depicted in Figure 1 below. The approach followed by Hospital B is summarily depicted in Figure 2 later. As an example that depicts the performance of the learned neural networks, the performance of the neural network for the 4^{th} clinical encounter, i.e., $m = 4$, is shown in Figure 3 in the Results Section.

$$\tanh(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}} \tag{2}$$

$$y_{m+1}^{(\hat{k})} - y_m^{(k)} = \sum_{j=1}^{r} w_j^{(m)} \tanh\left(\sum_{i=1}^{m} u_i^{(m)} x_i^{(k)} + v_i^{(m)} y_i^{(k)} + b_1^{(m)}\right) + b_2^{(m)} \tag{3}$$

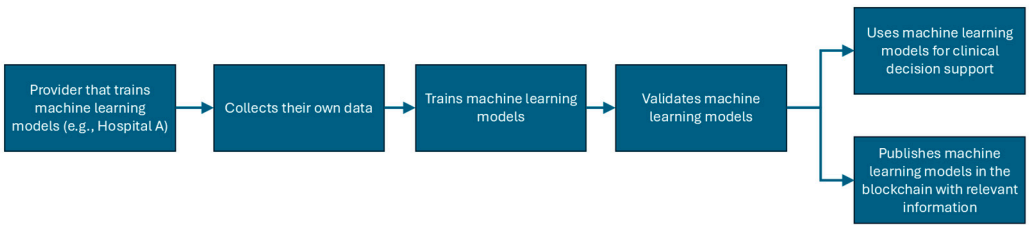


Figure 1. Approach followed by a healthcare provider that trains machine learning models (e.g., Hospital A).

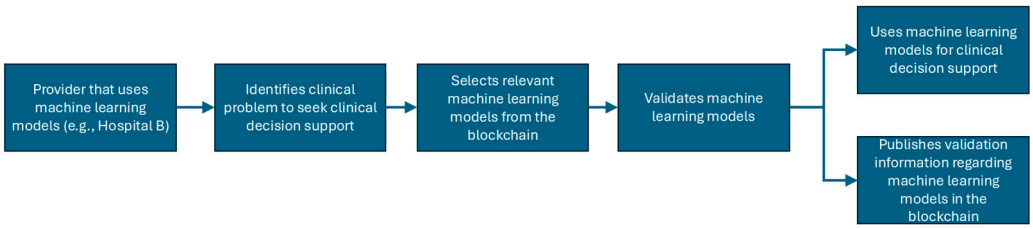


Figure 2. Approach followed by a healthcare provider that validates and uses machine learning models (e.g., Hospital B).

3.3. Federated Learning Being Realized Through Model Sharing

After learning the parameters for each clinical encounter, suppose Hospital A publishes those parameters as shown in Figure 1 for the benefit of other hospitals through a means

such as a dedicated healthcare blockchain. Now suppose another hospital (Hospital B) makes use of the models published by Hospital A.

Hospital B receives another cohort of similar cancer patients, the same number, i.e., $N = 385$. These patients are also governed by the same constraints and governing equations that were described for Hospital A. Hospital B also treats these patients with $x_1^{(k)} = 0.5, \forall k$ at the first clinical encounter. However, in the subsequent clinical encounters, the way of treatment is different. The cohort of patients in Hospital B receives treatment with the assistance of the machine learning models produced by Hospital A, by means of them being used as clinical decision support tools in Hospital B.

In Hospital B, the treatment intensity for a patient is calculated by solving the minimization problem shown in Equation (4). The objective here is to reduce the tumor radius of a patient k to some desired threshold denoted by y_{ref}, $y_{ref} \in \mathbb{R}^+$. For computation purposes, we set $y_{ref}, = 0.1$. The notation $\| \ \|$ in Equation (4) indicates the Euclidian norm.

$$x_m^{(k)*} = \arg\min_{x_m^{(k)}} \left\| y_{m+1}^{(\hat{k})} - y_{ref} \right\|^2 \tag{4}$$

The shared models are used in Hospital B with the data of a specific patient presenting at Hospital B. Decision support thus comes through a model that is making use of the data of a specific patient. Therefore, in the time of solving the optimization problem given in Equation (4), a model becomes a digital replica of a presenting patient—thereby the model becomes a digital twin of a patient at the time of solving the optimization problem.

As such, Hospital B derives some clinical decision support through the models generated by Hospital A thereby realizing the federated learning and model sharing approach we propose. The outcomes of these clinical interventions in Hospital B assisted by clinical decision support are shown in Figures 4 and 5. As shown in those figures, a significant improvement in clinical outcomes can be seen through this simulation. This demonstrates the feasibility and the likely impact of the proposed approach.

4. Results

The results of the conducted simulation are summarized in this section. In Figure 3, an example is shown, depicting the performance of a neural network learned by Hospital A; the neural network learned for clinical encounter 4 is shown. The correlation coefficients for training (50% of the data used), testing (25% of the data used), and validation (25% of the data used) stages are shown along with the overall performance.

In Figure 4, the patient outcomes are depicted. The progression of tumor radii along time, or the clinical encounters is plotted in Figure 4. In Figure 4a, the 385 patients from Hospital A that underwent the standard treatment protocol are plotted. In Figure 4b, the 385 patients from Hospital B that underwent treatment assisted by clinical decision support derived through the LNN models produced by Hospital A are plotted. The downward trend of the tumor progression of the patients from Hospital B is clearly visible in Figure 4b in contrast to Figure 4a, thereby depicting the potential clinical impact that our proposed approach of model sharing can have.

In Figure 5, we have depicted the relative reduction in tumor radius. We have plotted the ratio between the tumor radius of a patient at the first clinical encounter and the same at the last clinical encounter. That comparison also shows a significant reduction in tumor size among the patients from Hospital B in comparison to Hospital A. This is another indicator of the potential clinical impact our proposed approach of model sharing can have in improving patient outcomes.

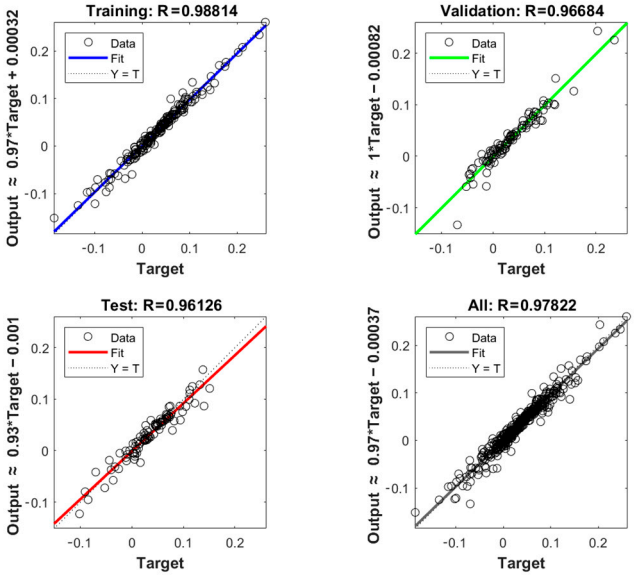


Figure 3. An example depicting the performance of a learned neural network; the one learned for clinical encounter 4.

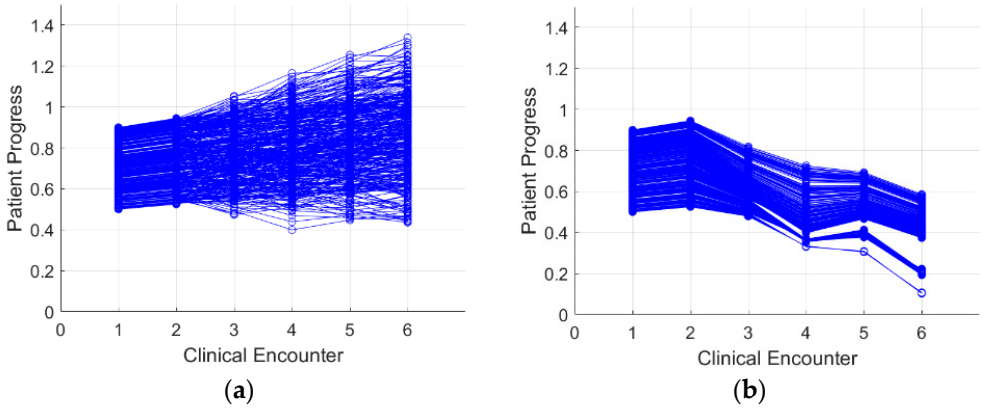


Figure 4. Progression of tumor radii (i.e., the variation in tumor radii): (**a**) a total of 385 patients from Hospital A that underwent the standard treatment protocol; (**b**) 385 patients from Hospital B that underwent treatment assisted by clinical decision support derived through LNN models produced by Hospital A.

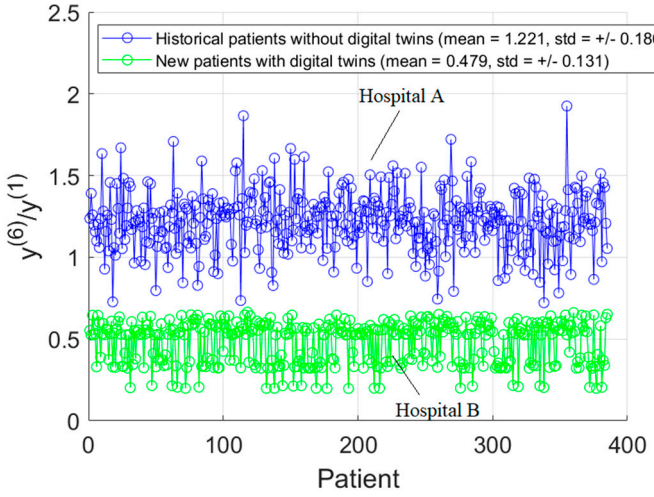


Figure 5. The impact of model sharing and intelligent clinical decision support: relative reduction of tumor size, i.e., $y_6^{(k)}/y_1^{(k)}$.

5. Model Sharing

5.1. A Blockchain Architecture for Model Sharing

For model sharing, we propose two blockchains: (1) model training blockchain, and (2) model validation blockchain. Shown in Table 1 is an example block of the model training blockchain showing the information that will be contained within the block. Shown in Table 2 is an example block of the model validation blockchain showing the information that will be contained within the block. The key is that the model training blockchain provides sufficient details for another subscribed healthcare provider to extract parameters and model architecture to use a machine learning model. The key with the validation blockchain is that it provides information about how well a model has been validated. Having information about the level of validation of a model helps other users to have more confidence in certain models over others. This helps in selecting suitable models as clinical decision support tools. This validation might have been conducted by the provider that has trained the model, or the validation might have been conducted by another provider, or several other providers. When several providers have validated, a model will have a validation history. This validation history further increases the indicative reliability of a model. In our approach, a key difference is that there is no model aggregation performed by a central server to converge onto a global model, as, in healthcare, this might carry risks in terms of liability. Instead, we allow the local learning of models relying on local data and evidence available to healthcare providers, and we share the learned models.

It is assumed that there will be a software layer that automates to some degree and validates these blockchain entries to ensure that manual entries are limited, and, thereby, errors in manual data entry are limited. This software layer that automates and validates the entries will help increase the trust among users about the accuracy of the information provided in the blockchains. The key is that, through these blockchains, the sensitive and raw health data are not shared, but learned models are shared with some details about them.

Table 1. An example block of the model training blockchain showing the information that will be contained within the block.

Field	Value
Targeted disease (e.g., cancer type and subtype)	C50 Breast cancer C50.1: Central portion of breast
Unique ID of the hospital that trained the model	#HospitalA#
The date at which the model was trained	20 December 2024
Unique ID for the model	#HospitalATrained20/12/2024ClinicalEncounter2TrainingSample385 TrainingAccuracy98%#
Clinical encounter (i.e., m = 2, 3, 4 or 5 with respect to the numerical example in Section 3)	$m = 2$
Number of neurons in the hidden layer (i.e., r)	$r = 10$
Model structure (LNN structure in Equation (3) as per the numerical example in Section 3)	$\hat{y_3^{(k)}} - y_2^{(k)} = \sum_{j=1}^{r} w_j^{(m)} \tanh\left(\sum_{i=1}^{m} u_i^{(m)} x_i^{(k)} + v_i^{(m)} y_i^{(k)} + b_1^{(m)} \right) + b_2^{(m)}$

Table 1. *Cont.*

Field	Value
Model parameters (i.e., the list of parameters and their values)	$w_1^{(m)} = \ldots,$ $w_2^{(m)} = \ldots,$ $\ldots,$ $\ldots,$ $w_r^{(m)} = \ldots,$ $u_1^{(m)} = \ldots,$ $u_2^{(m)} = \ldots,$ $\ldots,$ $\ldots,$ $u_m^{(m)} = \ldots,$ $v_1^{(m)} = \ldots,$ $v_2^{(m)} = \ldots,$ $\ldots,$ $\ldots,$ $v_m^{(m)} = \ldots,$ $b_1^{(m)} = \ldots,$ $b_2^{(m)} = \ldots,$
Size of training sample (i.e., the number of patients from which the model has been trained on)	385
Training accuracy (i.e., indicative score about the model accuracy; R: correlation coefficient in the case of the LNN structure presented in Section 3)	98%

Table 2. An example block of the model validation blockchain showing the information that will be contained within the block.

Field	Value
Unique ID of the Hospital that validated the model	#HospitalB#
The date at which the model was trained	20 December 2024
Unique ID for the Trained Model	#HospitalATrained20/12/2024ClinicalEncounter2TrainingSample385 TrainingAccuracy98%#
Unique ID for the Validated Model	#HospitalATrained20/12/2024HospitalBValidated20/12/2024Clinical Encounter2TrainingSample385TrainingAccuracy99%Validation Sample80ValidationAccuracy97%#
Size of validation sample (i.e., the number of patients from which the model has been validated on)	80
Validation accuracy (i.e., indicative score about the model accuracy; R: correlation coefficient in the case of the LNN structure presented in Section 3)	97%

5.2. Model Selecting for Clinical Decision Support

Model selection must be performed to select the best models to be used as clinical decision support tools. Model selection will be conducted by healthcare providers that have subscribed to the blockchains. It is assumed that model selection will also be conducted through an available software layer. It is assumed that this software facility available for model selection will be similar to a search engine, where searchers can create Boolean logic statements and collect models according to the coded logic. For example, Boolean logic statements such as the following one in Figure 6 will be possible.

SELECT
LNN Model(s)
FOR
C50 Breast Cancer
FOR
C50.1: Central portion of breast
FOR
Clinical Encounter 2
HAVING
Training Sample ≥ 300
AND
Training Accuracy ≥ 95%
AND
Validation Sample ≥ 50
AND
Validation Accuracy ≥ 90%

Figure 6. Example Boolean logic statement for selecting models from the blockchains.

The statements can be as complex or simple as need be. The healthcare providers will have the liberty to create their own logical statements and extract models as required for their purposes. If the providers train new models or validate existing models, there will be software facilities to push their developments to the blockchains so that other healthcare providers can access the latest developments. Executing logical statements as in Figure 6 will likely yield a set of models (i.e., several models that fit the logical criteria). In that case, it will be the extractor's duty to validate the cohort of models against the extractor's own data. This will in turn be a validation step that can be pushed back up to the blockchain. Through that validation, the extractor can select the statistically best models, and the validation they perform will be beneficial to other users as that will be new validation information regarding existing models. Thus, through this model sharing approach, federated learning and shared intelligence can be achieved.

6. Discussion

6.1. Summary of the Conducted Work

Answering the following research question was attempted in this paper: "How might the barrier of health data sharing be overcome through federated learning, model sharing, and digital twin-assisted clinical decision-making?" To answer this question, we proposed a novel approach that ensures the health data remain federated with the healthcare providers, for instance, with the hospitals, and each healthcare provider employs machine learning to train models on their own data so that these models can eventually function as clinical decision support tools. Then, instead of sharing the data, the models are shared via a private blockchain dedicated to healthcare providers. This private blockchain is made accessible to healthcare providers who have subscribed to it. This approach allows healthcare providers to access machine learning models produced by different healthcare providers and use them as clinical decision support tools.

To demonstrate the effectiveness of sharing machine learning models, a numerical example was presented in this paper. Synthetic data were used to demonstrate the application of Liquid Neural Networks (LNNs) to model cancer tumor radius. This paper showed how LNNs can be used as clinical decision support tools to plan treatments. This paper showed how hospitals can train such models using their own data and share these models, rather than the data itself, with other healthcare providers via a dedicated blockchain.

This paper also demonstrated how other hospitals can access and use these shared models as clinical decision support tools and how this ultimately improves patient outcomes. When using shared models for clinical decision making, they are used with the data of a particular patient. At this point, when leveraging data of a particular patient, the shared models become digital replicas of presenting patients. Thus, when functioning as clinical decision support tools, the shared models become digital twins of patients. This paper thus demonstrated the feasibility of using federated learning, model sharing, and the realization of digital twins, facilitated by blockchain, as an alternative to data sharing in healthcare, thereby answering the target research question.

6.2. Contribution to Theory

This paper contributed to the theory of LNNs by demonstrating their applicability in a healthcare setting targeted at chronic disease management. As a case example, modeling cancer tumor radius was attempted. The possibility to learn adapting neural networks that adapt according to patients and their clinical encounters was shown in this paper. By modeling tumor growth, the paper highlighted the potential of LNNs to support clinical decision making. The use of LNNs in this paper further demonstrated the capability of LNNs to handle dynamic data. Thereby the potential of LNNs to function as clinical decision support tools was demonstrated. This assists treatment planning and enhances the precision and personalization of care. By enhancing precision and personalization through the use of patient data at the point of functioning as clinical decision support systems, the LNN models function as digital twins of patients.

This paper also contributed to theory by demonstrating the impact of sharing machine learning models in healthcare. This paper demonstrated a means for hospitals to share learned machine learning models instead of raw data. An approach was laid out to make tailored use of blockchain technology to share machine learning models in such a way that other healthcare providers can access and use them for clinical decision support. The proposed approach addresses some of the privacy and security concerns associated with health data sharing. The proposed approach not only facilitates federated learning but also enhances clinical decision support through shared intelligence while ensuring that sensitive patient data remain protected.

6.3. Contribution to Practice

This paper contributed to practice by proposing an approach to tailor and use blockchain to share machine learning models among healthcare providers instead of sharing sensitive healthcare data. Healthcare providers can learn machine learning models from the federated data that they possess and then can securely share models, ensuring transparency and traceability, while raw health data are kept private. This method allows healthcare providers to benefit from shared intelligence through shared machine learning models used as clinical decision support tools, without the risks associated with having to share raw healthcare data. The proposed approach thereby assists the enhancement of clinical decision making and entailing care whilst preserving the privacy of data.

6.4. Limitations of This Study

One limitation of this paper can be noted to be this work being based on numerical examples produced on synthetic data. This work was carried out on synthetic data mainly because of two reasons: (1) the difficulty of accessing real-world health data; and (2) the insufficiency and incompleteness of real-world healthcare data even if they are accessible. Our purpose was not to show an implementation study that accesses real-world healthcare at this stage. Our purpose was rather to provide a proof of concept that sets a foundation for future developments in this very nascent area of the next generation digital transformation of healthcare. Building on the concepts we have presented here, future work can develop towards implementation.

7. Conclusions

This paper serves to demonstrate the feasibility of performing federated learning and sharing machine learning models among healthcare providers for the machine learning models to be used as clinical decision support tools. As a case example, learning Liquid Neural Networks (LNNs) through federated learning for clinical decision support and sharing the learned models through a tailored blockchain architecture was proposed.

The findings suggest that federated learning and model sharing can indeed enhance clinical decision making and improve the precision and personalization of treatment through the shared models functioning as digital twins of patients at the point of providing clinical decision support. The proposed approach thereby facilitates shared intelligence in healthcare along with intelligent data-driven clinical decision support whilst preserving the privacy of sensitive healthcare data. This in turn addresses several data security and privacy issues associated with health data sharing. This is one of the first papers to explore this approach in the healthcare context.

Future work will focus on implementation. Expanding further to include real-world hospitals and real-world data will be interesting. Not being limited to LNNs as we have done in this paper and exploring the use of other advanced machine learning models underpinned by the model sharing approach proposed in this paper will also be interesting.

With the introduction of these technological advancements to healthcare, investigating the long-term impacts on patient outcomes and healthcare delivery is also essential. Additionally, addressing specific concerns regarding health data sharing, such as data ownership, consent, intellectual property, and regulatory compliance, will also be crucial for the broader adoption of approaches such as the one presented in this paper.

Author Contributions: N.W. contributed through research leadership, conceptualization, writing and reviewing the article, and sourcing required funding. N.U. contributed through conceptualization, data generation, modeling, analysis, and writing and reviewing the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The generation of synthetic data used in this work, data analysis, and presentation were carried out using MATLAB R2022b. The data and MATLAB code are available upon request. Please contact Nilmini Wickramasinghe over n.wickramasinghe@latrobe.edu.au or nilmini.work@gmail.com, or Nalika Ulapane over n.ulapane@latrobe.edu.au or nalika.ulapane@gmail.com.

Acknowledgments: Optus and La Trobe University, Australia, are acknowledged for their support of this work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access* **2019**, *7*, 167653–167671. [CrossRef]
2. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2405–2415. [CrossRef]
3. Björnsson, B.; Borrebaeck, C.; Elander, N.; Gasslander, T.; Gawel, D.R.; Gustafsson, M.; Jörnsten, R.; Lee, E.J.; Li, X.; Lilja, S. Digital twins to personalize medicine. *Genome Med.* **2020**, *12*, 1–4. [CrossRef] [PubMed]
4. Ștefănigă, S.A.; Cordoș, A.A.; Ivascu, T.; Feier, C.V.I.; Muntean, C.; Stupinean, C.V.; Călinici, T.; Aluaș, M.; Bolboacă, S.D. Advancing Precision Oncology with Digital and Virtual Twins: A Scoping Review. *Cancers* **2024**, *16*, 3817. [CrossRef]
5. Wickramasinghe, N.; Ulapane, N.; Nguyen, T.A.; Andargoli, A.; Ossai, C.; Shuakat, N.; Zelcer, J. Towards discovering digital twins of dementia patients: Matching the phases of cognitive decline. *Alzheimer's Dement.* **2022**, *18*, e066336. [CrossRef]
6. Liu, Y.; Zhang, L.; Yang, Y.; Zhou, L.; Ren, L.; Wang, F.; Liu, R.; Pang, Z.; Deen, M.J. A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access* **2019**, *7*, 49088–49101. [CrossRef]
7. Wickramasinghe, N.; Ulapane, N.; Andargoli, A.; Shuakat, N.; Nguyen, T.; Zelcer, J.; Vaughan, S. Digital twin of patient in clinical workflow. *Proc. R. Soc. Vic.* **2023**, *135*, 72–80. [CrossRef]
8. Hölbl, M.; Kompara, M.; Kamišalić, A.; Nemec Zlatolas, L. A systematic review of the use of blockchain in healthcare. *Symmetry* **2018**, *10*, 470. [CrossRef]
9. McGhin, T.; Choo, K.-K.R.; Liu, C.Z.; He, D. Blockchain in healthcare applications: Research challenges and opportunities. *J. Netw. Comput. Appl.* **2019**, *135*, 62–75. [CrossRef]
10. Prokofieva, M.; Miah, S.J. Blockchain in healthcare. *Australas. J. Inf. Syst.* **2019**, *23*. [CrossRef]
11. Wickramasinghe, N.; Jayaraman, P.P.; Forkan, A.R.M.; Ulapane, N.; Kaul, R.; Vaughan, S.; Zelcer, J. A vision for leveraging the concept of digital twins to support the provision of personalized cancer care. *IEEE Internet Comput.* **2021**, *26*, 17–24. [CrossRef]
12. Ayoub, O.; Andreoletti, D.; Knapińska, A.; Goścień, R.; Lechowicz, P.; Leidi, T.; Giordano, S.; Rottondi, C.; Walkowiak, K. Liquid Neural Network-based Adaptive Learning vs. Incremental Learning for Link Load Prediction amid Concept Drift due to Network Failures. *arXiv* **2024**, arXiv:2404.05304.
13. Bidollahkhani, M.; Atasoy, F.; Abdellatef, H. LTC-SE: Expanding the potential of liquid time-constant neural networks for scalable AI and embedded systems. *arXiv* **2023**, arXiv:2304.08691.
14. Chahine, M.; Hasani, R.; Kao, P.; Ray, A.; Shubert, R.; Lechner, M.; Amini, A.; Rus, D. Robust flight navigation out of distribution with liquid neural networks. *Sci. Robot.* **2023**, *8*, eadc8892. [CrossRef] [PubMed]
15. Hasani, R.; Lechner, M.; Amini, A.; Rus, D.; Grosu, R. Liquid time-constant networks. In Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21), Virtually, 2–9 February 2021; pp. 7657–7666.
16. Um, K.; Hu, X.; Thuerey, N. Liquid splash modeling with neural networks. In *Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2018; pp. 171–182.
17. Banning, M. A review of clinical decision making: Models and current research. *J. Clin. Nurs.* **2008**, *17*, 187–195. [CrossRef]
18. Musen, M.A.; Middleton, B.; Greenes, R.A. Clinical decision-support systems. In *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 795–840.
19. Bica, I.; Alaa, A.M.; Lambert, C.; Van Der Schaar, M. From real-world patient data to individualized treatment effects using machine learning: Current and future methods to address underlying challenges. *Clin. Pharmacol. Ther.* **2021**, *109*, 87–100. [CrossRef]
20. Hlávka, J.P. Security, privacy, and information-sharing aspects of healthcare artificial intelligence. In *Artificial Intelligence in Healthcare*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 235–270.
21. Li, L.; Fan, Y.; Tse, M.; Lin, K.-Y. A review of applications in federated learning. *Comput. Ind. Eng.* **2020**, *149*, 106854. [CrossRef]
22. Dhade, P.; Shirke, P. Federated Learning for Healthcare: A Comprehensive Review. *Eng. Proc.* **2024**, *59*, 230. [CrossRef]
23. Prayitno; Shyu, C.-R.; Putra, K.T.; Chen, H.-C.; Tsai, Y.-Y.; Hossain, K.T.; Jiang, W.; Shae, Z.-Y. A systematic review of federated learning in the healthcare area: From the perspective of data properties and applications. *Appl. Sci.* **2021**, *11*, 11191. [CrossRef]
24. Gad, A.G.; Mosa, D.T.; Abualigah, L.; Abohany, A.A. Emerging trends in blockchain technology and applications: A review and outlook. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 6719–6742. [CrossRef]
25. Khezr, S.; Moniruzzaman, M.; Yassine, A.; Benlamri, R. Blockchain technology in healthcare: A comprehensive review and directions for future research. *Appl. Sci.* **2019**, *9*, 1736. [CrossRef]
26. Agbo, C.C.; Mahmoud, Q.H.; Eklund, J.M. Blockchain technology in healthcare: A systematic review. *Healthcare* **2019**, *7*, 56. [CrossRef]
27. Wickramasinghe, N.; Ulapane, N. Digital Twins of Patients: A Liquid Neural Network Interpretation. In Proceedings of the Australasian Conference on Information Systems, Canberra, Australia, 4–6 December 2024. Available online: https://aisel.aisnet.org/acis2024/32 (accessed on 5 February 2025).

28. Fosgate, G.T. Practical sample size calculations for surveillance and diagnostic investigations. *J. Vet. Diagn. Investig.* **2009**, *21*, 3–14. [CrossRef] [PubMed]
29. Rossi, A.; Chiodini, P.; Sun, J.-M.; O'Brien, M.E.; von Plessen, C.; Barata, F.; Park, K.; Popat, S.; Bergman, B.; Parente, B. Six versus fewer planned cycles of first-line platinum-based chemotherapy for non-small-cell lung cancer: A systematic review and meta-analysis of individual patient data. *Lancet Oncol.* **2014**, *15*, 1254–1262. [CrossRef] [PubMed]
30. Roy, S.K.; Manna, S.; Dubey, S.R.; Chaudhuri, B.B. LiSHT: Non-parametric linearly scaled hyperbolic tangent activation function for neural networks. In Proceedings of the International Conference On Computer Vision And Image Processing, Nagpur, India, 4–6 November 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 462–476.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.