

Article — Position doctrinale · Sur inscription

# Clinically-Informed Neural Networks (CINNs)

*Quand la littérature médicale devient contrainte d'apprentissage — et pourquoi ce n'est pas de la régularisation*

Jérôme Vetillard · March 2025 · Twingital Institute

## CE QU'IL FAUT RETENIR

- Les PINNs incorporent des lois physiques comme contraintes d'apprentissage. Les CINNs transposent ce principe au domaine clinique où seule la connaissance statistique populationnelle est disponible.
- $\mathcal{L} = \mathcal{L}_{\text{données}} + \lambda \cdot \mathcal{L}_{\text{clinique}}(f(x), \theta_{\text{lit}})$  — la littérature médicale publiée contraint l'espace des solutions admissibles.
- Ce n'est pas de la régularisation : la contrainte opère sur  $\mathcal{L}$  (trajectoires), pas  $\theta$  (paramètres). Elle est sémantiquement spécifiée par le domaine.
- Trois propriétés formelles admissibles : différentiabilité (P1), indépendance de source (P2), traçabilité épistémique (P3).
- Architecture originale Qualees — détails d'implémentation réservés pour publication ultérieure après dépôt de protection intellectuelle.

**Note préliminaire — Protection intellectuelle.** Les CINNs constituent une contribution architecturale originale en cours de protection intellectuelle. À ce titre, cet article se limite volontairement à l'exposition du cadre théorique, du positionnement épistémologique et du positionnement dans la littérature. Les détails d'implémentation ne seront pas abordés ici et feront l'objet d'une publication ultérieure après dépôt.

## 1. L'analogie inspiratrice : les PINNs

Les Physics-Informed Neural Networks ont résolu un problème élégant : comment incorporer une loi physique connue dans l'apprentissage d'un réseau de neurones, sans que cette loi soit une donnée d'entraînement supplémentaire, mais une contrainte sur l'espace des solutions admissibles.

La fonction de perte devient :

$$\mathcal{L} = \mathcal{L}_{\text{données}} + \lambda \cdot \mathcal{L}_{\text{physique}}$$

où  $\mathcal{L}_{\text{physique}}$  mesure la violation de l'équation différentielle connue. Le réseau apprend depuis les données et reste contraint par les lois de la physique. Il ne peut pas converger vers une solution numériquement satisfaisante mais physiquement absurde.

Raissi, Perdikaris et Karniadakis (2019) ont montré que cette architecture permet d'apprendre des dynamiques complexes avec des données remarquablement rares — précisément parce que la contrainte physique agit comme régularisateur structurel, réduisant dramatiquement l'espace des solutions à explorer.

Un point important souvent mal compris :  $\mathcal{L}_{\text{physique}}$  n'est pas évalué seulement aux points de données. Il est évalué sur un ensemble de collocation points disséminés dans le domaine. Ce qui permet à la contrainte physique d'agir partout, y compris là où aucune observation n'existe.

**La question centrale** : que se passe-t-il quand vous ne disposez pas de l'équation — mais que vous disposez d'une connaissance statistique dense sur le domaine ?

## 2. Le problème spécifique à la biologie clinique

En physique, les lois sont connues, déterministes, universelles. La thermodynamique d'une turbine est la même à Paris qu'à Tokyo. L'équation de Navier-Stokes ne dépend pas de l'histoire personnelle du fluide.

Des PINNs ont été appliqués avec succès à certains domaines biologiques — pharmacocinétique/pharmacodynamique, modélisation hémodynamique, électrophysiologie cardiaque. Ces applications fonctionnent précisément là où des équations mécanistiques partielles existent (Rackauckas et al., 2020 ; Ottens et al., 2022 ; Sahli Costabal et al., 2020).

Mais ces domaines représentent une fraction du spectre clinique. La dynamique d'une pathologie chronique complexe — progression de la maladie rénale chronique, trajectoire de l'insuffisance cardiaque, évolution d'une pathologie neurodégénérative — n'obéit pas à une équation différentielle connue et partagée.

Il n'existe pas d'équation différentielle exacte de la progression de l'insuffisance rénale chronique. Il existe des distributions de trajectoires, des corrélations entre biomarqueurs, des prévalences d'événements cliniques issues de décennies de recherche clinique publiée.

*Ce n'est pas de l'ignorance. C'est un régime épistémologique différent : non pas l'absence de connaissance, mais une connaissance de nature statistique et populationnelle, là où les PINNs supposent une connaissance déterministe et individuelle.*

À cela s'ajoute le problème structurel des données cliniques : le régime HDLSS — High Dimension, Low Sample Size. Quatre consultations par an. Des données éparses, hétérogènes, avec des mécanismes de données manquantes non aléatoires (MNAR). Dans ce régime, un réseau sans contrainte externe sur-apprend sur la cohorte d'entraînement et généralise mal précisément là où la généralisation est l'enjeu clinique (Sugiyama et al., 2012).

## 3. L'insight central : la littérature clinique comme source de contraintes

La réponse que nous explorons chez Qualees avec TweenMe repose sur une observation simple mais structurante.

La littérature médicale publiée — essais randomisés, études de cohorte, méta-analyses, registres — contient une information statistique considérable sur les trajectoires des pathologies modélisées. Cette information est encodée sous des formes variées : distributions de biomarqueurs, relations entre variables biologiques, prévalences et incidences d'événements cliniques, distributions de survie.

Cette information n'est pas dans les données d'entraînement du modèle. Elle est dans la connaissance collective de la communauté scientifique, distribuée dans des milliers de publications, construite sur des populations souvent bien plus larges et mieux caractérisées que la cohorte d'entraînement locale.

La fonction de perte prend la forme :

$$\mathcal{L} = \mathcal{L}_{\text{données}} + \lambda \cdot \mathcal{L}_{\text{clinique}}(f(x), \theta_{\text{lit}})$$

où  $\mathcal{L}_{\text{clinique}}$  mesure l'écart entre les sorties du modèle  $f(x)$  et les paramètres statistiques  $\theta_{\text{lit}}$  — distributions de biomarqueurs, intervalles de confiance d'événements cliniques, corrélations inter-variables documentées.

La contrainte  $\theta_{lit}$  est estimée depuis la littérature publiée, construite sur des cohortes distinctes de la cohorte d'entraînement locale. Une limite structurelle doit être reconnue : la littérature publiée et la cohorte d'entraînement locale peuvent partager des biais systématiques de sélection. L'indépendance de source ne garantit pas l'indépendance distributionnelle.

#### 4. Positionnement dans la littérature existante

Le positionnement couvre six familles de travaux : (4.1) knowledge distillation et teacher-student frameworks (Hinton et al., 2015), (4.2) constrained optimization in neural networks (Marquez-Neila et al., 2017 ; Pathak et al., 2015), (4.3) Bayesian deep learning avec priors informés (Gal & Ghahramani, 2016), (4.4) Universal Differential Equations et modèles hybrides QSP (Rackauckas et al., 2020 ; Gábor et al., 2021), (4.5) validation externe et TRIPOD (Collins et al., 2015 ; Wynants et al., 2020), (4.6) transfer learning et fine-tuning en ML clinique (Wornow et al., 2023).

Les CINNs occupent une position distincte : non pas là où les équations sont partiellement connues, mais là où seule la connaissance statistique populationnelle est disponible, sans structure différentielle sous-jacente. La source de connaissance n'est pas un modèle neuronal mais une base de connaissances statistiques traçable, révisable et auditable.

#### 5. Pourquoi ce n'est pas de la régularisation ?

La régularisation L1/L2 contraint l'espace des paramètres  $\tau$ . Les CINNs déplacent la contrainte vers l'espace des trajectoires  $\tau$  :

$$\tau_{CINN} = \tau_{données} + \lambda_{bio} \cdot \tau_{clinique}(\tau, \theta_{lit})$$

où  $\tau$  désigne la trajectoire prédite et  $\theta_{lit}$  les paramètres statistiques issus de la littérature.

*Une contrainte de régularisation dit : « préfère les solutions simples parmi toutes les solutions mathématiquement possibles. » Une contrainte CINN dit : « préfère les solutions biologiquement plausibles — en définissant la plausibilité biologique par ce que la communauté scientifique a établi empiriquement. »*

La différence avec la calibration classique : dans les CINNs, les contraintes sont estimées depuis une source externe indépendante. La calibration utilise les données pour calibrer le modèle sur les données. Les CINNs utilisent la connaissance externe pour contraindre le modèle là où les données sont insuffisantes.

Asymétrie critique : dans les PINNs, la contrainte est dure (équation différentielle). Dans les CINNs, elle est molle (distribution statistique). Un CINN contraint par une littérature biaisée échoue silencieusement — ce qui impose un monitoring continu en contexte MDR.

Trois propriétés formelles admissibles : **P1 — Différentiabilité** (contrainte active pendant l'entraînement), **P2 — Indépendance de source** (estimée depuis une source externe), **P3 — Traçabilité épistémique** (chaque paramètre traçable vers des sources primaires identifiées).

#### 6. Le problème du loss balancing

Wang, Teng et Perdikaris (2022) ont montré via l'analyse du Neural Tangent Kernel (NTK) que le déséquilibre des gradients entre les termes de perte est la principale source d'échec des PINNs entraînés naïvement. Ce phénomène est directement pertinent pour les CINNs.

Stratégies candidates : Self-Adaptive PINNs (SA-PINN, McClenny & Braga-Neto, 2023), gradient normalization (Chen et al., 2018), uncertainty weighting (Kendall et al., 2018).

Pour les CINNs, le problème de pondération présente une dimension supplémentaire : les  $\lambda$  doivent refléter non seulement l'équilibre numérique des gradients, mais aussi la valeur épistémique relative des sources. Une méta-analyse Cochrane devrait peser davantage qu'une étude rétrospective monocentrique.

## 7. Ce que les CINNs visent à apporter — et ce qu'ils ne résolvent pas

Bénéfices attendus sous réserve de validation : stabilité hors distribution en régime HDLSS, interprétabilité clinique (position relative dans les distributions de référence), traçabilité réglementaire (documentation technique MDR).

Limites structurelles non résolues : la qualité de l'extraction conditionne tout, le biais des sources est propagé (pas corrigé), la pondération des contraintes est un problème ouvert, la hiérarchie épistémique des sources (analogie GRADE) reste à formaliser, la temporalité des sources (versioning, obsolescence) est un composant de maintenance active.

## 8. Critères de validation et conditions de falsifiabilité

L'hypothèse centrale peut être énoncée de manière falsifiable : dans le régime HDLSS clinique, lorsque la cohorte d'entraînement locale est en deçà d'un seuil  $N$  *tel que l'estimation directe atteint une variance inacceptable, la contrainte externe apporte un gain de calibration mesurable — à condition que la divergence distributionnelle entre population de référence et cohorte locale soit en deçà d'un seuil  $\Delta$ .*

Cette hypothèse est rejetée si les expériences montrent systématiquement l'absence de gain ou une dégradation de la calibration sous contrainte.

## 9. Conclusion : la contrainte comme forme de connaissance

*Les données sont rares. La connaissance ne l'est pas.*

Des décennies de recherche clinique publiée ont produit une connaissance statistique considérable. Les CINNs sont une tentative de combler cette asymétrie — de faire de la connaissance collective une contrainte active dans l'apprentissage.

Dans la famille des architectures informées — PINNs pour la physique déterministe, Neural ODEs pour les dynamiques continues, UDE pour les systèmes partiellement connus — les CINNs proposent une position distincte : l'inférence sous contrainte épistémique clinique externe, pour des systèmes dont la dynamique est régie par des lois statistiques populationnelles plutôt que par des équations mécanistiques.

**Cette architecture est une hypothèse. Elle reste à valider.**

---

Thématiques : CINNs · PINNs · HDLSS · IA médicale · Contrainte clinique

---

Twingital Institute — [twingital-ventures.com](https://twingital-ventures.com) — Tous droits réservés

Ce document est la propriété de Jérôme Veillard / Twingital Institute. Toute reproduction, même partielle, est soumise à autorisation préalable.