PhysioMap: an ontology-grounded causal knowledge graph of human physiology

Robert Hoehndorf
King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
E-mail: robert.hoehndorf@kaust.edu.sa

Paul N. Schofield
Dept of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

Georgios V. Gkoutos
Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

Physiological mechanisms connect molecular changes to measurements such as blood pressure and plasma metabolite concentrations. We developed PhysioMap to represent these connections across biological scales. PhysioMap defines physiological traits through ontologies and distinguishes causal influence, production, constitution, quantitative identity, and modulation. A versioned projection maps entailed ontology patterns to typed constraints on quantitative structural causal models. A derivative-sign abstraction supports qualitative prediction of intervention responses at a stable equilibrium with feedback. The resource contains 1,699 traits and 2,387 typed relations. Expert review assessed 83 relations across all five types. In a rare metabolic disease evaluation against the Human Phenotype Ontology annotation corpus, the solver returned 175 directions across 870 downstream gene–phenotype pairs. Of these directions, 171 agreed with the reference, giving 97.7% agreement at 20.1% coverage. We compared forward predictions and inverse lesion rankings with shortest signed paths and signed diffusion on the same reference. PhysioMap connects physiological trait definitions to executable intervention queries and supports the assembly of models from explicitly typed physiological knowledge.

Keywords: systems biology; computational physiology; causal knowledge graph; ontology; structural causal model; qualitative abstraction; comparative statics; feedback; abduction.

1. Introduction

Physiological interventions can affect several organs through feedback. For example, inhibiting angiotensin-converting enzyme (ACE) reduces its activity in the renin–angiotensin–aldosterone system, which connects angiotensin II to vascular resistance, renal sodium handling, and blood pressure. A model must combine these dependencies to predict the direction of the blood-pressure response. A measured trait also needs a precise meaning: plasma angiotensin II concentration specifies a chemical, a quantity, and its compartment.

We define a physiological trait as a quality borne by an entity or process in a specified biological context. Forward prediction takes an intervention on one or more traits and computes which other traits increase or decrease after the system reaches a new equilibrium.

© 2026 The Authors. Open Access chapter published by World Scientific Publishing Company and distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License.

Abduction takes observed trait directions and ranks interventions that explain them. Rare metabolic diseases supply an application: an HFE-associated defect reduces hepcidin, which alters ferroportin-dependent iron handling and measurable iron traits. Here, a lesion variable denotes the primary activity or physiological function affected by a gene-associated disease mechanism.

Directed resources such as SIGNOR, OmniPath, INDRA, Reactome, and KEGG provide regulatory, signaling, and reaction knowledge.[1–5] Network propagation spreads information through their connections,[6,7] while sign-consistency and optimization methods find network explanations for measured molecular changes.[8,9] Predicting an intervention response in a feedback system additionally requires a rule for combining coupled mechanisms and opposing contributions. Parameterized differential equations supply that rule, and Physiome, CellML, BioModels, and HumMod provide reusable quantitative models.[10–13] Their assembly across organs requires compatible variables, equations, and parameters.

Earlier physiomaps represent physiological processes and participants, and Chalkboard traces qualitative perturbations through their networks.[14,15] Toonsi et al. connect knowledge graphs to causal and probabilistic semantics and use graph-guided mediation to analyze drug–disease effects in population data.[16] PhysioMap develops these ideas for quantitative physiological traits, relation-specific structural constraints, and qualitative equilibrium responses in cyclic systems.

Entity–quality ontology patterns define traits using anatomy, chemicals, proteins, processes, and qualities.[17–19] Web Ontology Language (OWL) reasoning classifies these definitions and entails relations among them. To compute interventions, these logical relations require an explicit causal interpretation.[20] Structural causal models (SCMs) supply such an interpretation: a perfect intervention replaces the equation for the intervened variable and retains the other equations.[21,22] Extensions to cyclic systems accommodate physiological feedback.[23,24] The representation must also distinguish causal influence from production, constitution, quantitative identity, and modulation, because these relations constrain different mathematical objects.

We collected physiological statements from textbooks, quantitative models, curated resources, and expert curation; classified their relation types; and encoded them in an OWL knowledge base. A versioned projection transforms entailed patterns into a typed causal knowledge graph (CKG) that constrains a class of quantitative SCMs. Derivative signs provide a qualitative abstraction of those constraints. We implemented a feedback-aware solver for first-order responses and a separate modulation query. We evaluate forward prediction of directional phenotypes and abduction of single-lesion hypotheses in rare metabolic disease, and use focused semantic checks to assess how the typed representation handles intervention.

Data and code availability. The knowledge base, projection registry, software, evaluation inputs, and expert review are available at `https://github.com/bio-ontology-research-group/physiomap/`. Supplementary material is hosted at `https://github.com/bio-ontology-research-group/physiomap/releases/download/v1.1.2/supplement.pdf`. An interactive map and interven-

tion interface is available at `https://bio2vec.net/physiomap/`. Software uses BSD 3-Clause; the map, benchmark data, and documentation use CC BY 4.0.

2. Methods

2.1. *Ontology and trait semantics*

We represent PhysioMap as an OWL 2 terminological knowledge base (TBox) $\mathcal{K}$ with standard OWL semantics.[25] Its upper level distinguishes continuants, processes, and traits. Imported modules provide classes for chemicals, proteins, cells, anatomy, processes, and qualities from Chemical Entities of Biological Interest (ChEBI), the Protein Ontology (PR), the Cell Ontology (CL), Uberon, the Gene Ontology (GO), and the ontology of phenotypic qualities (PATO).[17,26–30]

We use the entity–quality phenotype pattern to define a trait as a quality borne by an entity or process in a biological context.[18,19] A named trait T_i has the necessary conditions

$$\mathcal{K} \models T_i \sqsubseteq \mathsf{Trait} \sqcap \exists\, \mathsf{hasPart}.\big(E_i \sqcap \exists\, \kappa_i.C_i \sqcap \exists\, \mathsf{hasQuality}.Q_i^E\big), \qquad \kappa_i = \begin{cases} \mathsf{contextPartOf} & E_i \text{ continuant} \\ \mathsf{occursIn} & E_i \text{ process,} \end{cases} \tag{1}$$

where E_i is the characterized entity or process, C_i its recorded context, and Q_i^E an entity-specific subclass of the PATO quality Q_i. We omit the context conjunct when no context is recorded. The relation $\mathsf{contextPartOf}$ localizes a continuant, whereas $\mathsf{occursIn}$ localizes a process and propagates along parthood ($\mathsf{occursIn} \circ \mathsf{partOf} \sqsubseteq \mathsf{occursIn}$). The PATO hierarchy determines whether a quality requires a continuant or process bearer.

Parthood connects trait definitions across scales. For composable qualities, the reasoner groups traits by their bearers or their bearers' parts. For example, plasma volume, erythrocyte volume, and blood volume share the group *volume of blood or its parts*. Quantitative composition is specified separately through the identities and constitutive relations below. Extensive quantities such as volume admit these grouping rules; intensive quantities such as pressure require their own measurement model.

Causal axioms describe a type-level capacity. We encode that capacity with the collection pattern for biological function.[31] Each trait class T has an associated collection T^{all}, with $T \sqsubseteq \exists\, \mathsf{memberOf}.T^{\text{all}}$ and $T^{\text{all}} \sqsubseteq \exists\, \mathsf{hasMember}.T$. A causal axiom has the form $T_t^{\text{all}} \sqsubseteq \exists\, \mathsf{hasMember}.(\exists\, \mathsf{causedBy}.T_s)$: some member of the target collection is caused by an instance of the source. These existential axioms remain in the OWL 2 EL profile. Production and modulation use corresponding patterns, with modulation qualifying the causal influence whose slope it changes. Suppl. Sec. S1.2 gives the complete collection axioms. The projection supplies their interventional interpretation.

2.2. *Projection from ontology to a causal knowledge graph*

We next project the OWL knowledge base into a causal knowledge graph (CKG), that is, a knowledge graph with an explicitly interpreted causal edge type.[16] Let V be its named trait classes and $E_c \subseteq V \times V$ its causal-influence edges. The CKG also contains the typed production, constitution, quantitative-identity, and modulation relation instances defined below. A binary

relation instance $T_s \to T_t$ labeled r is present whenever $\mathcal{K}$ entails a relational pattern that the finite, versioned registry Π registers for relation type r. Each pattern $\rho \in \Pi$ defines an arity, direction, relation type, OWL class-expression premise φ_ρ, and the relation instance produced when that premise is entailed:

$$\mathcal{K} \models \varphi_\rho(T_s, T_t) \quad \Longrightarrow \quad T_s \xrightarrow{\rho} T_t. \tag{2}$$

Higher-arity patterns analogously produce typed hyperedges. The causal pattern, for example, has $\varphi_\rho(T_s, T_t) = \big(T_t^{\text{all}} \sqsubseteq \exists\, \mathsf{hasMember}.(\exists\, \mathsf{causedBy}.T_s)\big)$, so the entailment holding of ACE activity and angiotensin II yields the causal edge between those two traits.

Projection follows entailments.[32,33]

2.3. *Structural causal model semantics for typed relations*

The projection targets a class $\mathfrak{M}(\mathcal{K}, \Pi)$ of quantitative structural causal models. Trait T_i denotes an endogenous random variable X_i, and x_i a realized value. Exogenous random variables U_i enter the differentiable dynamics $\dot{x}_i = F_i(x_i, x_{\text{pa}(i)}, U_i)$. We assume a locally unique equilibrium x^* with a Hurwitz Jacobian and local self-regulation $\partial F_i / \partial x_i < 0$. A perfect intervention $\text{do}(x_A = \theta_A)$ replaces the mechanism or defining equation of each x_a by $x_a = \theta_a$, removes its incoming functional dependencies, and leaves all other equations unchanged.[21,34,35] Relation typing matters because it determines whether an intervention replaces a causal mechanism, a constitutive constraint, or a quantitative identity.

To define a direct effect, we let x_t re-equilibrate after an intervention on x_s while keeping every other variable at its reference value. Differentiating the target equilibrium equation and using negative self-regulation gives the direct-effect sign identity:

$$D_{st} := \left.\frac{dx_t^*}{d\theta}\right|_{\text{do}(x_s=\theta,\, x_j = x_j^*\ (j \neq s,t))} = -\,\frac{\partial F_t/\partial x_s}{\partial F_t/\partial x_t}, \qquad \text{sign}(D_{st}) = \text{sign}\left(\frac{\partial F_t}{\partial x_s}\right). \tag{3}$$

For a multivariate function, σ_j denotes the sign of its partial derivative with respect to argument j. Table 1 gives the quantitative constraint for each first-order relation type, its derivative-sign abstraction, and a PhysioMap example.

Table 1. Structural causal model semantics, derivative-sign abstractions, and examples for the first-order relation types. D_{st} is the direct effect in Eq. (3).

Relation	Quantitative constraint and derivative-sign abstraction	PhysioMap example
Causal influence	The target mechanism F_t depends non-constantly on x_s; $\sigma_{st} = \mathrm{sign}(\partial F_t/\partial x_s) = \mathrm{sign}(D_{st})$.	More plasma ACE directly raises plasma angiotensin II: $D_{\mathrm{ACE,AngII}} > 0$.
Production	$F_t = R_t + \epsilon_{pt} q_{pt}(x_p)$, with $q'_{pt} > 0$ and $\epsilon_{pt} \in \{-1, +1\}$; $\sigma_{pt} = \epsilon_{pt} = \mathrm{sign}(D_{pt})$.	A higher adrenal aldosterone-secretion rate raises plasma aldosterone.
Constitution	A whole is determined by its constituents: $X_w = C_w(X_{p_1}, \ldots, X_{p_r})$, with $\mathrm{sign}(\partial C_w/\partial x_{p_j}) = \sigma_j$ for $j = 1, \ldots, r$.	Adipose-tissue mass contributes positively to total body weight.
Quantitative identity	A named random variable is exactly a function of its arguments: $X_q = Q_q(X_{a_1}, \ldots, X_{a_\ell})$, with $\mathrm{sign}(\partial Q_q/\partial x_{a_j}) = \sigma_j$ for $j = 1, \ldots, \ell$.	Mean arterial pressure = cardiac output × total peripheral resistance; both partial derivatives are positive.

Sign annotations belong to a derived abstraction of the quantitative constraints. The abstract value $\sigma = ?$ denotes non-constant direct dependence with an unresolved derivative sign. Constitution and quantitative identity are directed: an intervention on the result replaces its defining equation while retaining the equations of its arguments.

A modulation $(m; (s, t), \mu)$, where μ is the mixed-derivative sign, is a (node, causal-influence edge) tuple in $\mathcal{H} \subseteq V \times E_c$. Its semantics is a paired intervention on modulator and source,

$$\widetilde{F}_t(\theta, \eta) := F_t\,|_{\mathrm{do}(x_s=\theta,\, x_m=\eta)}, \qquad \mathrm{sign}\left(\frac{\partial^2 \widetilde{F}_t}{\partial\eta\,\partial\theta}\right) = \mu. \tag{4}$$

With other arguments held fixed, the sign states whether increasing x_m makes the mechanism slope $\partial F_t/\partial x_s$ more or less positive. For example, cortisol increases the positive slope of the peripheral-resistance mechanism with respect to norepinephrine. Modulation is retained in $\mathcal{H}$ and supports a separate gain query. The first-order experiments use causal influence, production, quantitative identity, and constitution. A first-order effect of a modulator enters through an independently authored causal influence. The gain query combines the modulator's response with the mixed-derivative sign to determine how the signed slope changes. Suppl. Secs. S9.2 and S9.3 specify the second-order and population interpretations.

2.4. *First-order derivative-sign abstraction and feedback-aware prediction*

The quantitative SCM supplies the semantics; the PhysioMap solver reads only its first-order derivative-sign abstraction. The direct effects above keep the rest of the system fixed. Forward prediction instead asks for the total effect after every other variable re-equilibrates. At the

assumed equilibrium $F(x^*, \theta) = 0$, the total effect is the comparative static

$$\frac{dx^*}{d\theta} = -\left(\frac{\partial F}{\partial x}\right)^{-1} \frac{\partial F}{\partial \theta}. \tag{5}$$

The inverse Jacobian couples all variables in a feedback component. The signs of its entries may determine a response direction, or unavailable derivative magnitudes may leave it unresolved.[23,24,36] The PhysioMap solver returns `?` in the latter case.

The relation types constrain the SCM differently. The PhysioMap solver projects the first-order derivative constraints for causal influence, production, and quantitative identity into a signed dependency graph $G_\partial = (V, E_\partial)$, whose signed adjacency pattern defines one signed Jacobian. The PhysioMap solver represents constitution through a separate constitutive dependency graph $G_\kappa = (V, C)$ on the same vertices, with $(p_j, w) \in C$ labeled σ_j for each constituent in Table 1; adipose-tissue mass and lean-body mass each constitute total body weight. The constitutive dependency graph is acyclic because scales are strictly hierarchical, so no trait constitutes one of its own constituents, and determination is transferred along it in topological order by

$$\operatorname{sign}\left(\frac{dx_w^*}{d\theta}\right) = \bigoplus_{(p_j, w) \in C} \sigma_j \cdot \operatorname{sign}\left(\frac{dx_{p_j}^*}{d\theta}\right), \tag{6}$$

where $\cdot$ and $\bigoplus$ are the sign product and parallel combination defined below, and a constituted variable that also carries a causal sign combines the two the same way. Determination proceeds from constituents to the whole. The signed dependency graph can contain cycles, so computing downstream effects must account for feedback.

After intervention, we remove the incoming dependencies of the target from G_∂, yielding $G_\partial^{\mathrm{do}(k)}$, and decompose this graph into strongly connected components and process their acyclic condensation graph in topological order. A singleton uses sign algebra over $\{+, -, 0, ?\}$ with $+ \cdot + = - \cdot - = +$, $+ \cdot - = -$, and 0 annihilating products. An unknown factor absorbs a non-zero factor. The sign of a chain is the product of its edge signs. For parallel contributions, 0 is neutral, agreeing non-zero signs retain that sign, and opposing or unknown contributions give $?$. For a feedback component S, let J_S be its Jacobian sign pattern and b_S its signed forcing. The implicit function theorem and Cramer's rule give

$$\operatorname{sign}\left(\frac{dx_i^*}{d\theta}\right) = \operatorname{sign}\left(\frac{\det J_S^{(i)}}{\det J_S}\right), \qquad J_S^{(i)} = J_S \text{ with column } i \text{ replaced by } -b_S. \tag{7}$$

For a component with m variables, differentiating its equilibrium equations gives $J_S z = -b_S$. Cramer's rule gives Eq. (7). A real Hurwitz matrix has determinant sign $(-1)^m$; real eigenvalues contribute negative factors and conjugate pairs contribute positive products. Sound sign addition and multiplication ensure that a definite numerator sign holds for every compatible stable realization. Applying these operations in component order extends the argument to upstream forcing.[37–39] Suppl. Sec. S12.2 gives the full proof.

The exact component algorithm handles up to 16 traits, bounding each memoized determinant expansion by 2^{16} subproblems. Larger components use a conservative fixed-point

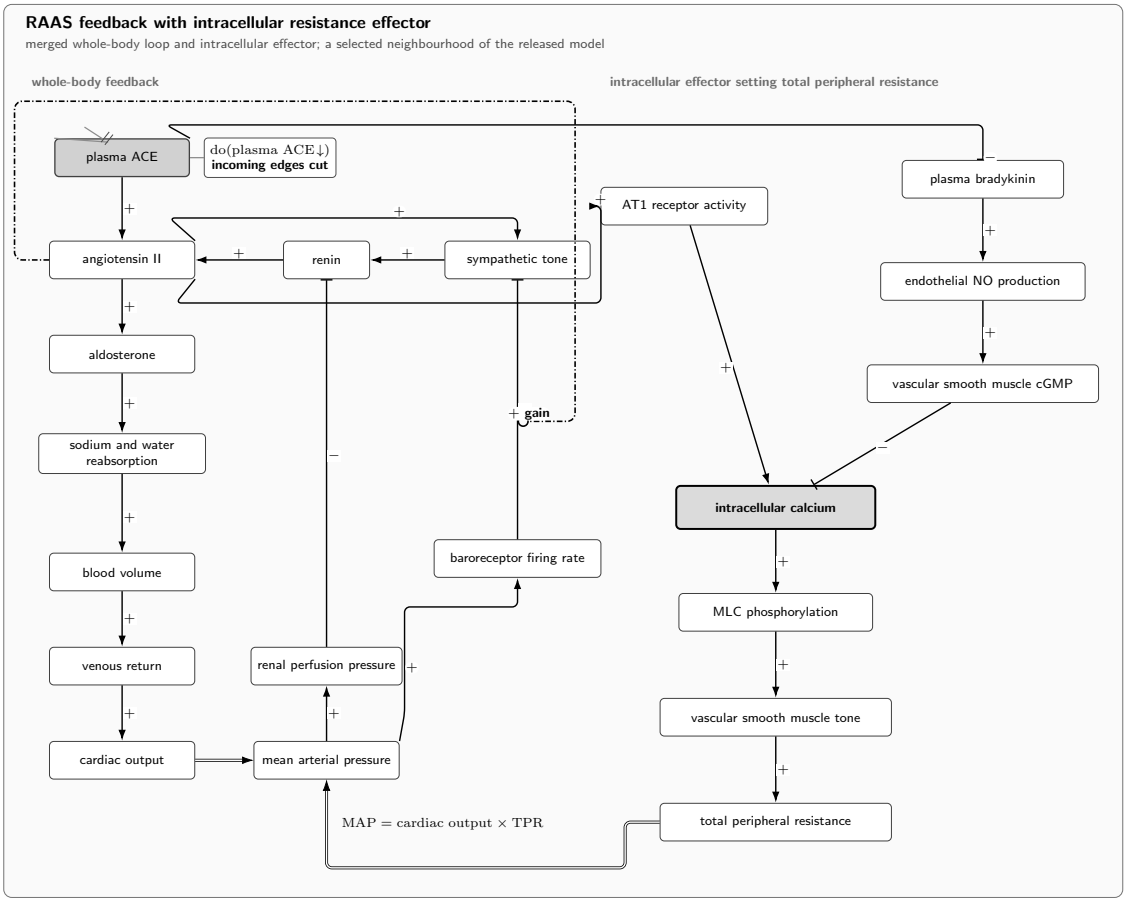


Fig. 1. A selected PhysioMap neighbourhood merges whole-body renin–angiotensin–aldosterone system (RAAS) feedback with an intracellular resistance effector. Pointed arrows are positive causal influences; bars are negative influences. These glyphs show the corresponding SCM derivative signs. The double line denotes the quantitative identity mean arterial pressure (MAP) = cardiac output × total peripheral resistance (TPR), and the dash-dot line denotes modulation of the inhibitory baroreceptor-to-sympathetic influence. The intervention on plasma angiotensin-converting enzyme removes incoming causal influences. AT1, angiotensin II type 1 receptor; NO, nitric oxide; cGMP, cyclic guanosine monophosphate; MLC, myosin light chain.

algorithm. Both algorithms return ? when their sign calculation remains unresolved. The guarantee is conditional on a locally unique stable equilibrium and negative self-regulation; the software takes these as model assumptions.

Figure 1 shows ACE inhibition perturbing volume and vascular-resistance pathways coupled by renal and baroreflex feedback. An intracellular arm connects receptor signals to smooth-muscle tone; typed relations distinguish causal influences, an identity, and modulation.

2.5. *Construction and content validation*

We assembled PhysioMap by direct expert curation, extraction from quantitative models and curated databases, and checked language-model proposals from textbooks. The cardiovascular

core used 17 Guyton BioModels modules and CellML integrators. Available equations supplied derivatives, while directional sources supplied constraints without complete functions. SIGNOR, the Adverse Outcome Pathway Wiki (AOP-Wiki), and direct literature curation provided additional content. Suppl. Sec. S15.1 lists the sources and acquisition metadata.

Following the entity–quality convention,[17–19] each candidate trait received an entity identifier, PATO determinable, optional context, stable label, and scale. We resolved identifiers against pinned releases, classified measurement kind as extensive, ratio, rate, or intensive, checked that bearer and quality categories matched, and reconciled duplicates. Only composable measurement kinds admit part-to-whole rules.[40,41] Unresolved traits remained explicitly primitive.

We assigned each relation to one of the five semantic types. Release checks covered schema, references, identifiers, bearers, constitution, relation typing, and regression against a frozen behavioral baseline. These checks assess formal consistency and preservation of expected behavior.

For cross-type content validation, P.N.S. reviewed a fixed-seed stratified sample of 83 v1.1.1 relations covering all five types. We interpreted `TRUE` as accepted, `FALSE?` as flagged for further review or investigation, and `FALSE` as rejected. We report the judgments for each sampled relation type. Suppl. Sec. S15.3 gives the protocol, sampling strata, and archive.

2.6. *Evidence and provenance model*

Evidence and provenance govern admission and document construction. Type, arguments, and context determine a relation's semantics, and its sign constrains a derivative. The solver reads relation types, arguments, contexts, and derivative signs.

For causal-influence candidates, the evidence model admits six classes. Perturbation, pharmacological intervention, genetic loss or gain of function, and Mendelian randomization form the interventional group because they use a manipulation or valid instrument. Mechanistic-model evidence derives a sign from a curated model, while curated-mechanistic evidence records a curated functional or mechanistic account; these form the mechanistic group. Admission requires evidence of the stated intervention or functional mechanism. Other relation types use type-specific evidence, and support awaiting assignment to a controlled evidence class remains unclassified. Each stable relation identifier links to support and a projection trace.

2.7. *Evaluation of prediction and abduction*

We evaluated forward prediction against the Human Phenotype Ontology (HPO) and its gene-to-phenotype annotation corpus (HPOA) from release 2026-02-16. For each gene entering the evaluation, we specified one or more primary lesion variables from gene-product identity and the disease mechanism. Suppl. Sec. S15.4 gives the mapping rule and links the complete versioned specification.

HPO phenotype classes follow an entity–quality (EQ) pattern, in which the entity E identifies the affected entity or process and the quality Q is drawn from PATO.[17,42] This pattern lets us identify a directional HPO class when Q states an increase or decrease relative to normal.

For example, hypoglycemia pairs blood glucose with decreased concentration (PATO:0001163). We encode the PATO distinction between increased quality (PATO:0002300) and decreased quality (PATO:0002301) as $s = \texttt{+}$ and $s = \texttt{−}$, respectively.

We propagated HPOA annotations only upward through the HPO subclass hierarchy. Let A_g be the HPO classes directly annotated to gene g, let HPO denote the HPO ontology, and let $M(a) = (X_i, s)$ map a directional HPO class a to PhysioMap variable X_i and direction s. The reference directions for X_i were

$$R_g(X_i) = \{s \in \{\texttt{+}, \texttt{−}\} \mid \exists h \in A_g \; \exists a : \mathsf{HPO} \models h \sqsubseteq a \;\wedge\; M(a) = (X_i, s)\}. \tag{8}$$

Therefore, a specific annotation h contributed the mapping of an ancestor a, never that of a descendant. We retained a gene–variable direction only when $R_g(X_i)$ was a singleton. We blocked propagation for two deoxycortisol terms whose asserted HPO ancestry would otherwise map them to cortisol, although neither term denotes cortisol (Suppl. Sec. S15.4). We did not score the intervened variable.

We scored all unambiguous downstream reference directions. To assess literature overlap between HPOA and PhysioMap, we repeated the analysis after excluding the 15 genes named in phenotype-directed curation fragments. Suppl. Sec. S15.4 gives the mapping sources and the overlap specification.

For each lesion, we applied the PhysioMap solver once and compared every downstream sign with the fixed reference pairs. Let C, W, and A denote correct, wrong, and abstaining pairs; missing and `?` predictions both count in A. We report precision $= C/(C + W)$ and coverage $= (C + W)/(C + W + A)$.

We compared two alternative inference rules on the same axioms and interventions. These rules differ only in how they combine directions, so the comparison evaluates the inference rules on shared physiological content. The shortest-path rule receives a lesion and an endpoint, selects one shortest directed path, and multiplies its edge signs. Signed diffusion receives a lesion and assigns a score to every downstream trait. Let S be the signed, column-normalized adjacency matrix, so a node with d explicitly signed outgoing arcs contributes $+1/d$ or $-1/d$ to each successor. We omit unknown-sign arcs before computing reachability and normalizing weights. Let e encode the signed intervention. We solved

$$r = (1 - \alpha)(I - \alpha S)^{-1} e, \tag{9}$$

taking $\mathrm{sign}(r_i)$ as the prediction and $|r_i|$ as its confidence. We set $\alpha = 0.85$ and swept every distinct confidence observed on the evaluation pairs to produce the precision–coverage curve.

We tested whether abstention concentrates shortest-path errors by permuting abstention labels within genes, conditional on each gene's path-error and abstention counts among path-answered pairs. The statistic counts errors assigned to abstained pairs; we report its exact one-sided upper-tail probability (Suppl. Sec. S15.4).

For the relation-typing ablation, we cumulatively added production, quantitative identity, and constitution to causal influence and repeated both tasks. Modulation uses a separate query. We checked typed intervention semantics with small models whose expected responses follow directly from their equations. The checks cover part-to-whole determination, production, quantitative identities, and the separate modulation query. Suppl. Sec. S15.7 gives the

models, expectations, and executable checks.

For abduction, we precomputed determinate phenotype signs for a fixed closed pool of single-lesion hypotheses and compared them with each observed directional profile. Agreement added one, contradiction subtracted one, and abstention scored zero. We ranked by net agreement, then fewer contradictions, then more agreements. Rank is one plus the number of strictly better candidates, so top-k, median rank, and mean reciprocal rank use the best position within a tie; unique top-1 requires no tie. We excluded the intervened trait and retained the true lesion in the pool.

Experiments used PhysioMap v1.1.1 (analysis v1.1.2). Suppl. Sec. S15.4 gives input checksums and file names; Suppl. Sec. S15.6 gives the reproduction commands.

3. Results

3.1. *Construction, projection, and expert validation of PhysioMap*

We built PhysioMap v1.1.1 as an OWL knowledge base containing 1,699 contextualized physiological traits and 2,270 causal-influence, 85 production, 4 constitution, 9 quantitative-identity, and 19 modulation axioms. Projection preserves evidence and provenance (Section 2.6).

Projection produced 2,387 typed relation instances as edges or hyperedges; the largest strongly connected component contains 213 traits.

The cross-type expert review accepted 69 of 83 sampled relations, flagged 12 for further investigation, and rejected 2 (Suppl. Table S1). The flagged category denotes an unresolved review judgment; both rejected relations were causal. These judgments identify content for correction or further investigation.

3.2. *Forward prediction of directional endophenotypes*

We tested how physiological dependencies recover the direction of measurements associated with rare metabolic disease. The HPO mapping provided downstream directions for 167 of 183 matched genes; the other 16 had a mapped direction only for the intervened variable. Across 870 downstream pairs, the solver returned 175 determinate directions: 171 agreed with the reference and 4 disagreed. Agreement was 97.7% at 20.1% coverage (Figure 2b).

Shortest paths returned 379 directions, with 292 agreements and 87 disagreements (77.0% agreement). Figure 2b compares signed diffusion across thresholds on the same reference and interventions. The overlap analysis retained 155 determinate predictions, of which 151 agreed with the reference.

A blinded audit of 60 sampled concordant pairs supported 57 reference directions and judged three context-dependent; all four discordant pairs supported PhysioMap (Suppl. Sec. S15.5). The primary result remains unadjusted at 97.7% agreement and 20.1% coverage.

PhysioMap and shortest paths returned identical directions on the 175 shared answered pairs, with 2.3% error. On the 204 additional pairs answered by shortest paths, PhysioMap abstained and the path error rate was 40.7%. The within-gene test gave $P = 2.05 \times 10^{-7}$: abstention concentrated cases difficult for path inference, explaining the precision–coverage trade-off.

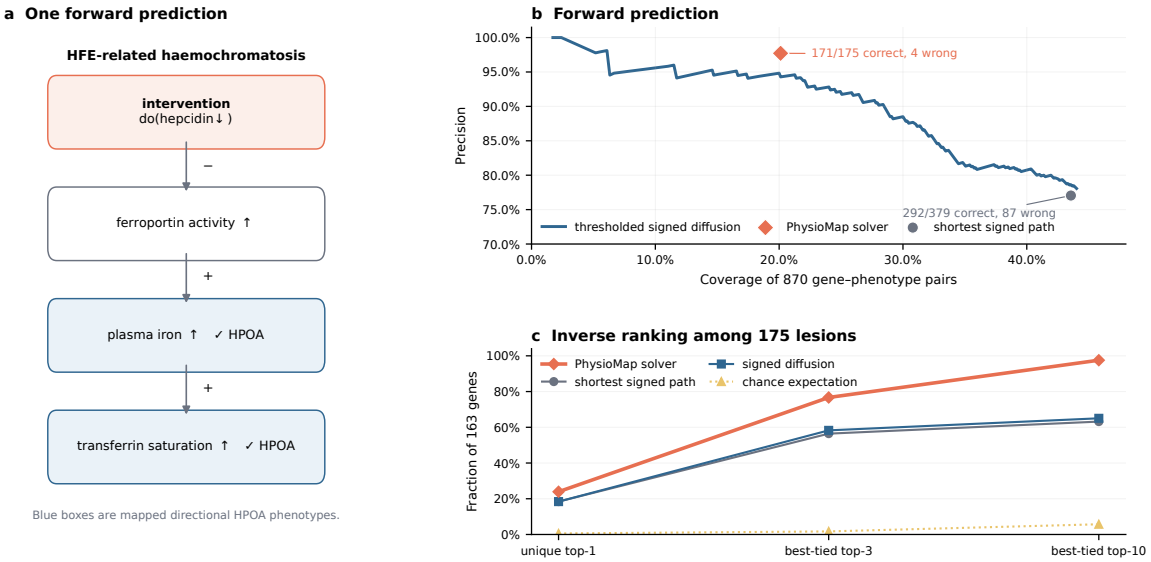


Fig. 2. Evaluation of prediction and abduction. (a) Relations supporting an HFE-associated hereditary haemochromatosis prediction; only downstream phenotypes enter scoring. (b) Precision versus coverage for thresholded signed diffusion, with operating points and exact counts for the PhysioMap solver and shortest signed path. (c) Abductive ranking within 175 known lesions for the three inference rules and chance expectation. Top-3 and top-10 use the best rank within ties, and a unique top-1 requires no tied candidate; mean reciprocal rank for the PhysioMap solver is 0.698.

3.3. *Abduction ranks candidate lesions from phenotype profiles*

The inverse benchmark excludes ARSA, CTNS, KYNU, and PKLR, which each have two primary lesion variables, leaving 163 genes. Each true lesion belongs to the closed pool of 175 hypotheses. Only the inference rule varies across comparisons (Table 2).

Table 2. Inverse lesion ranking over 175 candidates for 163 genes. Scoring, tie handling, and metrics are identical in every row; only the rule producing a candidate's predicted profile differs. Top-3 and top-10 use the best rank within ties, and a unique top-1 requires no tied candidate.

inference rule	unique top-1	top-3	top-10	MRR
PhysioMap solver	39	125	159	0.698
shortest signed path	30	92	103	0.528
signed diffusion	30	95	106	0.545
chance	1	3	9	0.033

The solver ranked the true lesion within the best-tied top three for 125 genes and uniquely first for 39 (Table 2).

Adding the first-order relation types gave no aggregate improvement in forward coverage or inverse ranking (Suppl. Table S3). The semantic checks verify typed intervention behavior

in small models; broader non-causal content would test its predictive contribution.

4. Discussion

PhysioMap's ontology defines physiological variables through shared entities, qualities, and contexts, providing a basis for interoperability across models and data sources. The same entity–quality pattern connects these variables to phenotype ontologies:[18,19] we used it to map directional HPO phenotypes to physiological changes. This connection allows phenotype annotations to inform both evaluation of predicted responses and inference of candidate lesions. The typed projection then gives the relations between variables their intervention semantics.

The solver extends qualitative inference in earlier physiomaps and Chalkboard[14,15] to responses coupled by cycles and feedback. Its derivative-sign abstraction supports predictions around a locally unique stable equilibrium with negative self-regulation. Opposing contributions can leave responses dependent on unknown derivative magnitudes. Quantitative systems-biology models can supply these magnitudes, resolve feedback responses, and assess stability numerically. PhysioMap could support model assembly from Physiome, CellML, and BioModels[10–12] by identifying shared variables and constraints on their dependencies; integration requires compatible units, contexts, and parameters. Regenerative mechanisms and transitions between equilibria require explicit dynamics, while coupled second-order modulation responses require further models and experiments.

Current forward coverage is 20.1% and depends on represented physiology and lesion mapping. Abstention concentrated cases with higher shortest-path error. PhysioMap and HPOA can share literature; the overlap analysis assesses phenotype-directed curation, while independent perturbation data would test transfer. Inverse ranking uses best-tied positions in a closed pool containing the true lesion; prospective open-set evaluation would test lesions outside that pool. Broader non-causal content would test the predictive contribution of the additional relation types. Content review used one co-author and unequal sampling rates; independent representative review would estimate release-wide quality. The reference audit concealed predictions and earlier judgments, but its co-author reviewer had previously assessed the four discordant pairs. Independent review would test reproducibility.

PhysioMap already contains a feedback component of 213 traits. Exact sign inference is restricted to small components; conservative inference on larger components leaves additional responses unresolved. Extending the map across human physiology will increase both curation demands and the complexity of dependencies across organs and scales. Reusable organ models could support expansion, but their shared variables and feedback require reconciliation. Ontology alignment can help identify these connections; unresolved context-specific conflicts currently yield unknown signs. Expansion should therefore be evaluated through predictive coverage, content quality, and computational cost.

By connecting phenotype observations to physiological mechanisms, PhysioMap provides a basis for reasoning about interventions and could guide the assembly of quantitative models across organs and scales.

References

1. L. Licata, P. Lo Surdo, M. Iannuccelli, A. Palma, E. Micarelli, L. Perfetto, D. Peluso, A. Calderone, L. Castagnoli and G. Cesareni, SIGNOR 2.0, the SIGnaling network open resource 2.0: 2019 update, *Nucleic Acids Research* **48**, D504 (2020).
2. D. Türei, A. Valdeolivas, L. Gul, N. Palacio-Escat, M. Klein, O. Ivanova, M. Ölbei, A. Gábor, F. Theis, D. Módos *et al.*, Integrated intra- and intercellular signaling knowledge for multicellular omics analysis, *Molecular Systems Biology* **17** (2021).
3. B. M. Gyori, J. A. Bachman, K. Subramanian, J. L. Muhlich, L. Galescu and P. K. Sorger, From word models to executable models of signaling networks using automated assembly, *Molecular Systems Biology* **13** (2017).
4. M. Milacic, D. Beavers, P. Conley, C. Gong, M. Gillespie, J. Griss, R. Haw, B. Jassal, L. Matthews, B. May *et al.*, The Reactome pathway knowledgebase 2024, *Nucleic Acids Research* **52**, p. D672–D678 (2024).
5. M. Kanehisa, M. Furumichi, Y. Sato, Y. Matsuura and M. Ishiguro-Watanabe, KEGG: biological systems database as a model of the real world, *Nucleic Acids Research* **53**, p. D672–D677 (2025).
6. A.-L. Barabási, N. Gulbahce and J. Loscalzo, Network medicine: a network-based approach to human disease, *Nature Reviews Genetics* **12**, p. 56–68 (2011).
7. L. Cowen, T. Ideker, B. J. Raphael and R. Sharan, Network propagation: a universal amplifier of genetic associations, *Nature Reviews Genetics* **18**, p. 551–562 (2017).
8. M. Gebser, T. Schaub, S. Thiele and P. Veber, Detecting inconsistencies in large biological networks with answer set programming, *Theory and Practice of Logic Programming* **11**, p. 323–360 (2011).
9. A. Liu, P. Trairatphisan, E. Gjerga, A. Didangelos, J. Barratt and J. Saez-Rodriguez, From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL, *npj Systems Biology and Applications* **5** (2019).
10. P. J. Hunter and T. K. Borg, Integration from proteins to organs: the Physiome project, *Nature Reviews Molecular Cell Biology* **4**, p. 237–243 (2003).
11. C. M. Lloyd, J. R. Lawson, P. J. Hunter and P. F. Nielsen, The CellML model repository, *Bioinformatics* **24**, p. 2122–2123 (2008).
12. R. S. Malik-Sheriff, M. Glont, T. V. N. Nguyen, K. Tiwari, M. G. Roberts, A. Xavier, M. T. Vu, J. Men, M. Maire, S. Kananathan *et al.*, BioModels: 15 years of sharing computational models in life science, *Nucleic Acids Research* **48**, D407 (2020).
13. R. L. Hester, A. J. Brown, L. Husband, R. Iliescu, D. Pruett, R. Summers and T. G. Coleman, HumMod: A modeling environment for the simulation of integrative human physiology, *Frontiers in Physiology* **2** (2011).
14. D. L. Cook, M. L. Neal, R. Hoehndorf, G. V. Gkoutos and J. H. Gennari, Representing physiological processes and their participants with PhysioMaps, *Journal of Biomedical Semantics* **4**, p. S2 (2013).
15. D. L. Cook, J. C. Wiley and J. H. Gennari, Chalkboard: ontology-based pathway modeling and qualitative inference of disease mechanisms, in *Pacific Symposium on Biocomputing 2007*, (World Scientific, 2007).
16. S. Toonsi, P. N. Schofield and R. Hoehndorf, Causal knowledge graph analysis identifies adverse drug effects, *Bioinformatics* **42** (2026).
17. G. V. Gkoutos, E. C. Green, A.-M. Mallon, J. M. Hancock and D. Davidson, Using ontologies to describe mouse phenotypes, *Genome Biology* **6** (2005).
18. R. Hoehndorf, A. Oellrich and D. Rebholz-Schuhmann, Interoperability between phenotype and anatomy ontologies, *Bioinformatics* **26**, 3112 (2010).
19. G. V. Gkoutos, P. N. Schofield and R. Hoehndorf, The anatomy of phenotype ontologies: prin-

ciples, properties and applications, *Briefings in Bioinformatics* **19**, p. 1008–1021 (2018).
20. A. Rector, R. Stevens and N. Drummond, What causes pneumonia? the case for a standard semantics for "may" in OWL, in *Proceedings of the 5th OWLED Workshop on OWL: Experiences and Directions*, , CEUR Workshop Proceedings Vol. 432 (RWTH Aachen University, 2008).
21. J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, 2009).
22. J. Woodward, *Making Things Happen: A Theory of Causal Explanation* (Oxford University Press, 2003).
23. J. M. Mooij, D. Janzing and B. Schölkopf, From ordinary differential equations to structural causal models: the deterministic case, in *Proceedings of the 29th Conference on Uncertainty in Artificial Intelligence (UAI)*, (AUAI Press, 2013). arXiv:1304.7920.
24. S. Bongers, P. Forré, J. Peters and J. M. Mooij, Foundations of structural causal models with cycles and latent variables, *The Annals of Statistics* **49** (2021).
25. B. Motik, P. F. Patel-Schneider and B. Cuenca Grau, OWL 2 Web Ontology Language: Direct Semantics (Second Edition) W3C Recommendation, (2012).
26. J. Hastings, G. Owen, A. Dekker, M. Ennis, N. Kale, V. Muthukrishnan, S. Turner, N. Swainston, P. Mendes and C. Steinbeck, ChEBI in 2016: Improved services and an expanding collection of metabolites, *Nucleic Acids Research* **44**, D1214 (2016).
27. C. J. Mungall, C. Torniai, G. V. Gkoutos, S. E. Lewis and M. A. Haendel, Uberon, an integrative multi-species anatomy ontology, *Genome Biology* **13**, p. R5 (2012).
28. A. D. Diehl, T. F. Meehan, Y. M. Bradford, M. H. Brush, W. M. Dahdul, D. S. Dougall, Y. He, D. Osumi-Sutherland, A. Ruttenberg, S. Sarntivijai, C. E. Van Slyke, N. A. Vasilevsky, M. A. Haendel, J. A. Blake and C. J. Mungall, The Cell Ontology 2016: enhanced content, modularization, and ontology interoperability, *Journal of Biomedical Semantics* **7**, p. 44 (2016).
29. The Gene Ontology Consortium, Gene ontology: tool for the unification of biology, *Nature Genetics* **25**, 25 (2000).
30. D. A. Natale, C. N. Arighi, J. A. Blake, C. J. Bult, K. R. Christie, J. Cowart, P. D'Eustachio, A. D. Diehl, H. J. Drabkin, O. Helfer *et al.*, Protein ontology: a controlled structured network of protein entities, *Nucleic Acids Research* **42**, D415 (2014).
31. R. Hoehndorf, L. Mencel, G. V. Gkoutos and P. N. Schofield, Large-scale reasoning over functions in biomedical ontologies, in *Formal Ontology in Information Systems: Proceedings of the 9th International Conference (FOIS 2016)*, , Frontiers in Artificial Intelligence and Applications Vol. 283 (IOS Press, 2016).
32. R. Hoehndorf, A. Oellrich, M. Dumontier, J. Kelso, D. Rebholz-Schuhmann and H. Herre, Relations as patterns: bridging the gap between OBO and OWL, *BMC Bioinformatics* **11**, p. 441 (2010).
33. M. A. Rodríguez-García and R. Hoehndorf, Inferring ontology graph structures using OWL reasoning, *BMC Bioinformatics* **19**, p. 7 (2018).
34. S. Beckers and J. Y. Halpern, Abstracting causal models, in *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, (AAAI Press, 2019).
35. P. K. Rubenstein, S. Weichwald, S. Bongers, J. M. Mooij, D. Janzing, M. Grosse-Wentrup and B. Schölkopf, Causal consistency of structural equation models, in *Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence (UAI)*, (AUAI Press, 2017). arXiv:1707.00819.
36. P. A. Samuelson, *Foundations of Economic Analysis* (Harvard University Press, Cambridge, MA, 1947).
37. K. Lancaster, The scope of qualitative economics, *The Review of Economic Studies* **29**, p. 99 (1962).
38. J. Quirk and R. Ruppert, Qualitative economics and the stability of equilibrium, *The Review of Economic Studies* **32**, p. 311 (1965).
39. J. Maybee and J. Quirk, Qualitative problems in matrix theory, *SIAM Review* **11**, p. 30–51

(1969).
40. P. Suppes, A set of independent axioms for extensive quantities, *Portugaliae Mathematica* **10**, 163 (1951).
41. D. H. Krantz, R. D. Luce, P. Suppes and A. Tversky, *Foundations of Measurement, Volume I: Additive and Polynomial Representations* (Academic Press, New York, 1971).
42. S. Köhler, M. Gargano, N. Matentzoglu, L. C. Carmody, D. Lewis-Smith, N. A. Vasilevsky, D. Danis, G. Balagura, G. Baynam, A. M. Brower *et al.*, The human phenotype ontology in 2021, *Nucleic Acids Research* **49**, p. D1207–D1217 (2021).