Interface Focus

Requirements for the formal representation of pathophysiology mechanisms by clinicians

de Bono, B.[1†,π], Helvensteijn, M.[Ω], Kokash, N.[Ω], Martorelli, I.[Ω], Sarwar, D.[†], Islam, S.[Ψ], Grenon, P.[π], Hunter, P.[†‡]

[†]Auckland Bioengineering Institute (ABI), University of Auckland, New Zealand
[π]Farr Institute, University College London, 222, Euston Road,United Kingdom
[Ω]Leiden Institute for Advances Computer Science, University of Leiden, The Netherlands
[Ψ]University of East London, University Way, London, United Kingdom
[‡]Department of Physiology, Anatomy and Genetics (DPAG), University of Oxford, United Kingdom

Abstract

Knowledge of multiscale mechanisms in pathophysiology is the bedrock of clinical practice. If quantitative methods, predicting patient-specific behavior of these pathophysiology mechanisms, are to be brought to bear on clinical decision-making, the Human Physiome community and Clinical community must share a common computational blueprint for pathophysiology mechanisms. A number of obstacles stand in the way of this sharing - not least the technical and operational challenges that must be overcome to ensure that (i) the explicit biological meaning of the Physiome's quantitative methods to represent mechanisms are open to articulation, verification and study by clinicians, and that (ii) clinicians are given the tools and training to explicitly express disease manifestations in direct contribution to modeling. To this end, the Physiome and Clinical communities must co-develop a common computational toolkit, based on this blueprint, to bridge the representation of knowledge of pathophysiology mechanisms (a) that is implicitly depicted in electronic health records and the literature, with (b) that found in mathematical models explicitly describing mechanisms. In particular, this paper makes use of a step-wise description of a specific disease mechanism as a means to elicit the requirements of representing pathophysiological meaning explicitly. The computational blueprint developed from these requirements addresses the Clinical community goals to (i) organize and manage healthcare resources in terms of relevant disease-related knowledge of mechanisms, and (ii) train the next generation of physicians in the application of quantitative methods relevant to their research and practice.

[1] Correspondence to: b.bono@ucl.ac.uk

Introduction

Practicing physicians and surgeons leverage knowledge of pathophysiology mechanisms to take clinical decisions. While the physiological community[2] generated this knowledge as a result of precise quantitative analysis, the pathophysiology-based methods employed by the clinical community for the training and practice of medicine primarily rely on qualitative approaches (given, for instance, the heavy reliance on free text descriptions and static diagrams in both pathophysiology textbooks, as well as in clinical record-taking). This divergence in approaches between the Physiome and Clinical communities is at the root of considerable loss of opportunity to share knowledge and collaborate on the study of pathophysiology mechanisms relevant to:

1) understanding disease processes, in terms of cause and consequence,
2) developing and re-purposing drugs and devices, as well as
3) planning diagnostic and therapeutic interventions.

To allay the collaborative handicap caused by the above discrepancy, it is crucial to:

A. identify the common ground – *i.e.* the knowledge scenarios in pathophysiology that the two communities have in common,
B. understand the implications of the lack of formal convergence between the knowledge management strategies of these two communities, and
C. provide solutions that improve cross-disciplinary convergence in pathophysiology knowledge management.

In addressing points A to C above, this paper identifies and collates the representational requirements for the exchange of computer-readable knowledge about pathophysiology mechanisms between the Physiome and Clinical communities. This blueprint of requirements builds a shared view of pathophysiology mechanisms – a view that is consistent and compatible with the operational goals of both communities. In the conclusion section, we make recommendations for this blueprint as the basis of an open computational toolkit for the management of pathophysiology resources in tangible support of cross-disciplinary collaboration.

[2] The Physiome community is distinct from the physiological community. The Physiome community has developed modelling and data standards, and associated computational tools and repositories, to create a reproducible multiscale physiological modelling framework. We will refer to both the physiological community and the Physiome community from here on as the 'Physiome' community.

Point A: The common ground
Knowledge of multiscale anatomy is leveraged by both Physiome and Clinical communities to organize knowledge about the:

1) location of biomedical measurement, and
2) routes linking locations of measurement.

Both communities develop biophysical models, to a varying degree of formality, that take into account correlations between biomedical measurements. In some cases, it is important to understand how the route linking the locations of measurement contributes to the correlation. In addition, mathematical depictions of biophysical mechanisms, in both normal physiology and pathophysiology, may be represented in terms of rates of transfer over these routes.

Furthermore, both communities take into account the creation, destruction and alteration of routes as leveraged by the body to regulate rates of transfer over a wide range of physical dynamic systems, for instance in:

1) **Normal Physiology**
    i. **immune system** (*e.g.* the puncturing of cytosol-to-extracellular connections by the complement cascade to eliminate virus-infected and cancer cells, the repair of blood vessel leakage by coagulation),
    ii. **cell signalling and metabolism** (*e.g.* endocrine and exocrine cells fuse internal vesicles to plasma membrane to transfer secretory substances to the outside of the cell),
    iii. **pumps/channels** (*e.g.* connectivity route of urinary bladder to the outside is reversibly altered by pelvic sphincters for the purpose of waste elimination; voltage-gated of ion channels across the phospholipid bilayers),
    iv. **development** (*e.g.* the growth of the septum transversum – precursor to the diaphragm - that establishes a topological separation between the thoracic and abdominal cavities; the fusion of left and right endocardial tubes to form the primitive heart; intussusceptive angiogenesis in the creation of new blood vessels).

2) **Pathophysiology**

i. **infection** (*e.g.* filariasis infestation leading to blockage of lymphatic channels, helicobacter pylori causing gastric ulcer perforation leading to abnormal communication between the stomach lumen and blood vessel lumen),

ii. **metabolic accumulation** (*e.g.* atherosclerosis leading to mycardial infarction due to coronary artery blockage or artic valve regurgitation, and gallstone formation leading to acute pancreatitis due to blockage at the Ampulla of Vater, excessive communication between the dermis and the outside environment due to keratin/keratinocyte-adhesion weakness in epidermolysis bullosa and pemphigus vulgaris respectively; aortic valve incompetence due to atherosclerosis causing regurgitation between the left ventricle and the aorta),

iii. **external trauma** (*e.g.* axonotmesis of the median nerve due to carpal tunnel syndrome, intestinal obstruction due to incarcerated hernia, fractured ribs causing an abnormal communication between the pleural space and the outside air in traumatic pneumothorax),

iv. **toxicity** (*e.g.* tetrodotoxin poisoning leading to muscle paralysis due to blockage of fast voltage-gated sodium channels, cholera toxin results in diarrhoea due to massive salt and water secretion across the intestinal epithelium),

v. **congenital malformations** (*e.g.* obstruction of the foramina of Luschka leading to hydrocephalus due to decrease in ventricular cerebrospinal fluid outflow, post-partum patent ductus arteriosus and atrial septal defect leading to abnormal pulmonary-artery-to-aorta and left-atrium-to-right-atrium communication respectively).

3) **Therapeutic Intervention**

A number of therapeutic modalities seek to rectify disease conditions through the manipulation of anatomical routes. A number of drugs, for instance, directly affect membrane channels (*e.g.* Amiodarone blocking both sodium and potassium channels in the myocardium, omeprazole inhibiting the proton pump in the epithelium of the stomach) or the patency of vessels (*e.g.* adrenergic decongestants reduce respiratory mucosal hyperaemic

swelling and mucus production). In addition, there is much in the practice of surgery that consists of topological operations at a macroscopic level (*e.g.* a Roux-en-Y anastomosis that creates end-to-side route between stomach and jejunum or ileum, repair of gastric ulcer, coronary artery bypass graft).

The common ground between Physiome and Clinical communities for the representation of pathophysiology mechanisms, therefore, consists of the application of topological concepts to describe routes of transfer across multiscale anatomical compartments. Two of the key obstacles for interoperability between the two communities are the lack of:

1. a formal and shared topological reference map over which representations of mechanisms by Physiome and Clinical communities can be linked, together with
2. agreed methods and communal tools to describe the location of measurement in terms of this map.

The implications of these obstacles are discussed in the next section.

Point B: Lack of convergence

Considerable effort has been invested by:

a) the Physiome community to ensure that the mathematical description of pathophysiology process correlations is formalized, standardized and shareable (*e.g.* models in CellML (1) and SBML (2)), and
b) the Clinical community to develop shareable topographical models of anatomy for the interpretation of histological and radiological data (*e.g.* (3)(4)), as well as shareable data models for healthcare data management (*e.g.* (5)).

However, the above investment has a somewhat limited effect on the consistent bridging of [i] the physiological meaning of mathematical models of processes to [ii] the physiological meaning of histological, radiological and healthcare models.

Multiscale anatomical knowledge of measurement location, and routes linking such locations, is not formally described and shareable between the two communities. Developing a formal topological representation of biological structure across multiple

scales to address the above shortcoming is a considerable challenge. While, for instance, substantial investment by the Systems Biology community has gone into the clear cataloguing of well-characterized compendia of relevant biological structures (*i.e.* proteins (6), small molecules (7) and subcellular locations (8)) that participate in molecular mechanisms, the unmanageable size of the combinatorial space bearing every possible anatomical compartment has severely curbed the ability of any single structural model of multiscale anatomy to enumerate them all. The challenge for the Physiome and Clinical communities remains, therefore, to provide the means to systematically map the anatomical location of a wide range of measurements. The following scenarios indicate some of the shortcomings to be surmounted:

a) **Clinical**:

i. **Pharmaceutical – linking tissue data to pharmacokinetic models**
The pharmacokinetic (PK) modeling of processes of drug absorption, distribution, metabolism and elimination (ADME) is a core component in drug discovery and development. In physiology-based PK modeling (PBPK), account is taken of the key anatomical sites responsible for ADME functions, as well as the routes by which these sites communicate. Some examples of hand-drawn schematics showing key PBPK variables and rates of transfer over relevant anatomical compartments are depicted in **Figure 1**. A key bottleneck in the development of PBPK models is sourcing of data about measurements of gene expression (*e.g.* (9)) and drug permeability (*e.g.* (10)(11)) for tissues, relevant to the modeling of (i) drug effect/toxicity, and (ii) blood-to-tissue drug transitions in ADME.
Shortcoming: there is currently no provision for the systematic and formal recording of anatomical location relevant to ADME PBPK model variables, gene expression or drug permeability.

ii. **Surgical – management challenges in patients with diabetes mellitus**
Diabetes mellitus (DM) complications interfere directly with drug ADME. Specifically, relevant complications include the following departures from normal ADME function (12)(13):

- Absorption: post-prandial blood pooling, stasis of stomach and gallbladder, intestinal hurry due to oversecretion or bacterial overgrowth in the small intestine, constipation,
- Distribution: redistribution of blood volume, altered drug-to-plasma protein binding, and changed volumes of drug distribution,
- Metabolism: the ability of the liver to produce plasma protein (*e.g.* albumin) and to metabolize drugs is impaired, as well as
- Elimination: impaired biliary and renal function that impact on drug biotransformation and excretion respectively.

Shortcoming: the lack of a coherent representation of the routes over which ADME processes are hampered in DM is a key reason why pharmaceutical development of drugs does not take into account the wide alteration in ADME that typifies patients with DM. DM patients, therefore, are vulnerable to therapeutic errors and drug interactions because of this shortcoming.

b) **Physiome - bridging data and model variables**

The Physiome community have developed repositories to manage the storage and serving of models and data (*e.g.* (14)). However finding data that is relevant to a particular model is still a considerable challenge. For example, data from anatomical sites relevant to blood pressure regulation need to be semantically matched to the variables in mechanistic models of hemodynamic control (*e.g.* (15)(16)). In particular, the challenge is to link:

i. data such as blood pressure in the lumen of the segment of internal carotid artery containing the carotid sinus (**Table 1.**, measurement A), rates of secretion of adrenalin from chromaffin cells in the adrenal medulla (measurement B), and rates of secretion of renin from granular cells in the juxtaglomerular apparatus in the kidney (measurement C), and
ii. equation-linked variables from the Guyton model of circulation control (17)(18) as depicted in **Figure 2** and further explained in Table 1.

Both data and model variables are a type of measurement that is ascribed to a particular anatomical location. Identifying a functional link between the above

data and variables, therefore, requires the automated comparison of location knowledge for the measurements (*i.e.* data and variables) outlined in Table 1. A corresponding schematic that attempts to relate the locations of these entities in terms of anatomical routes relevant to the mechanisms of blood pressure control is shown in **Figure 3**.

Shortcomings:

a. the locations listed in Table 1 are described in free-text and, as such, do not provide not computer-readable knowledge supporting the inferencing of anatomical routes bridging these sites. The quality of these location descriptions, while very high to the human reader, is of limited utility to the automated management and integrative goals of an electronic repository for physiology modeling resources.
b. similarly, while it was possible to manually develop Figure 2 to suggest the functional relationship between measurements and model variables, the means to automatically generate such circuitboard diagrams of anatomical routes given the locations in Table 1 as input, to our knowledge, does not exist.

To allay the above shortcomings, the next section focuses on the step-wise depiction of a specific pathophysiology mechanism to elicit the representational requirements for the recording of (i) the location of measurements, and (ii) the routes over which processes responsible for the correlation of these measurements unfold.

Point C: Requirements for pathophysiology knowledge management

This first part of this section describes the biological steps in the pathophysiology mechanism of hydronephrosis caused by calculus obstruction of the ureter. In the second part, this biological description is converted into a graph linking the locations of measurements, relevant to the evolution of this pathology, over an anatomical route that accounts for the correlation between measurements. Generalized representational requirements for a database of pathophysiology pathways are discussed in the third part.

1) The pathophysiology of hydronephrosis

There are a number of conditions (19) that lead to urinary solutes precipitating out of solution to create calculi, for instance:

a. Dehydration, which increases the concentration of all solutes due to the kidneys' attempts to reduce the amount of water lost to urine;
b. Secretory: certain urinary membrane channels may oversecrete a lithogenic ion (*e.g.* calcium, oxalate) or undersecrete an ion that keeps lithogenic ions in solution (*e.g.* magnesium, citrate);
c. Sustained metabolic changes: *e.g.* altered diet or therapy (such as cancer chemotherapy, loop diuretics, carbonic anhydrase inhibitors) that increase the production of lithogenic solutes (*e.g.* urate) that have to be eliminated by the kidney;
d. Inflammation (*e.g.* due to infection or autoimmune insult), which causes mucus proteoglycans to be excessively secreted into urine. The presence of a stone further increases epithelial irritation/inflammation and overlying infection;
e. Strictures to urinary outflow (*e.g.* scarring of epithelial tube due to inflammation) leading to urinary stasis (*i.e.* more time for a solute to precipitate out of solution) and bacterial overgrowth (*i.e.* more mucus secretion);

The calculus formation, therefore, increases the likelihood of further stone accretion by stimulating conditions #d and #e above. The growth of a stone in the pelvicalyceal region of the urinary tract provides the right conditions for a calculus to reach to a size that cannot be subsequently conveyed down the ureter (20). Such an accretion becomes lodged at the pelviureteric junction, reducing urinary outflow. This flow reduction is followed by a build up of upstream pressure as the nephrons in that kidney continue to produce urine. This urinary pressure build up goes on to compress the blood supply of the kidney in the hilum, leading to vascular strangulation and subsequent atrophy of that organ.

2) Formal representation of cause and effect in the hydronephrosis scenario

The above informal account of the pathophysiology mechanism of hydronephrosis implicitly describes a number of correlations of rate and state measurements drawn

from a range of locations along and across the renal epithelial (*i.e.* urinary tract) and endothelial (*i.e.* blood vessel) conduit systems. Therefore, we look to the basic geometric configuration of these conduit systems to motivate the organization of structural knowledge about the location of these measurements. The basic organizational features that need to be taken into account to represent routes linking measurement locations are that:

iii. key biophysical interactions regulating the constitution, as well as the flow, of fluid in the lumen take place between the Wall (W) of the conduit and the Content (C) contained by the same conduit – two basic types of location, therefore, need to be distinguished: (C, W);

iv. transfers between conduits systems transit through the connective tissue 'glue' that organizes bundles of endothelial, epithelial and neural conduits;

v. the representation of long range material transfers (*i.e.* beyond the range of the diffusion limit) must take into account the topology of conduit arborisations to explicitly link tube types from different material properties that communicate along the same conduit system (*e.g.* ipsilateral nephrons and ureter).

The type of biomedically relevant measurement in the above scenario is of two kinds: the state property of some Material (M) or the rate property of some Process (P). At the very least, therefore, there are four distinct types of located measurement entities, symbolized as follows: P_W, P_C, M_W, M_C.

In this pathophysiology scenario, correlations of located measurements can be biophysically modeled over the six equations below (numbered I to VI). These functional relationships define transfers taking place within C, within W or across CW, as follows:

I. The changing composition of urine; precipitation of salts

The concentration of various biochemical substances in urine, such as calcium, urate, phosphate, *etc* ($\boldsymbol{C_i}$, i=1,2…), changes with time at rates…

$$\frac{d\boldsymbol{C_i}}{dt} = f_i\{\boldsymbol{C_1}, \boldsymbol{C_2}, \boldsymbol{C_3}, \}$$

…dependent on various biochemical processes. At some point an ion such as calcium (concentration $\boldsymbol{C_1}$) supersaturates and precipitates out of solution, at which point a kidney stone begins to form.

II. Formation of calcified material

The mass of calcified material, $\boldsymbol{m_{Ca}}$, accumulates at a rate…

$$\frac{dm_{Ca}}{dt} = g_1\left\{-\frac{dC_1}{dt}, \boldsymbol{I}, ..\right\}$$

…dependent on the rate of decrease of calcium in solution ($-\frac{dC_1}{dt}$) and other factors such as inflammation ($\boldsymbol{I}$, e.g. due to infection or irritation), which causes mucus proteoglycans to be excessively secreted into urine.

III. Changing flow in the ureter

Flow in the ureter, $\boldsymbol{q_{ureter}}$, depends on the diameter of the ureter at the point of stone formation ($\boldsymbol{d_{ureter}}$), the mass of the kidney stone ($\boldsymbol{m_{Ca}}$), the filtrate flow rate ($\boldsymbol{q_{filtrate}}$) and the viscosity of urine ($\boldsymbol{\nu}$):

$$\boldsymbol{q_{ureter}} = g_2\{\boldsymbol{d_{ureter}}, \boldsymbol{m_{Ca}}, \boldsymbol{q_{filtrate}}, \boldsymbol{\nu}\}$$

If the stone grows large enough to block the ureter, the flow will stop, but the consequences of stone formation is felt upstream well before this point is reached.

IV. Distension of the renal calyces and the renal pelvis

The difference between filtrate flow ($\boldsymbol{q_{filtrate}}$) and flow in the ureter downstream of the stone ($\boldsymbol{q_{ureter}}$) is absorbed by distension of the renal calyces and the renal pelvis (which have a total volume $\boldsymbol{V_r}$):

$$\frac{dV_r}{dt} = \boldsymbol{q_{filtrate}} - \boldsymbol{q_{ureter}}$$

The fluid pressure in this space (p_r) depends on both the volume $\boldsymbol{V_r}$ and the elasticity $\boldsymbol{E_t}$ of the surrounding tissue:

$$\boldsymbol{p_r} = g_3\{\boldsymbol{V_r}, \boldsymbol{E_t}\}$$

V. Pressure in the surrounding tissue

The pressure $\boldsymbol{p_t}$ in the surrounding tissue depends on the pressure $\boldsymbol{p_r}$ in the renal calyces and the renal pelvis and on the elasticity $\boldsymbol{E_t}$ of the tissue:

$$\boldsymbol{p_t} = g_4\{\boldsymbol{p_r}, \boldsymbol{E_t}\}$$

VI. Compression of the renal vasculature

The increasing pressure in the surrounding tissue $\boldsymbol{p_t}$ compresses the renal vasculature and increases the resistance to arterial blood flow $\boldsymbol{q_a}$, which also depends upon geometric and material properties $\boldsymbol{E_a}$ of the vascular wall and the pressure $\boldsymbol{p_a}$ in the blood:

$$\boldsymbol{q_a} = g_5\{\boldsymbol{p_t}, \boldsymbol{E_a}, \boldsymbol{p_a}\}$$

The above relations provide the topology of a graph of C- and W-located M and P measurements depicting the key types of transfer relevant to the pathology of hydronephrosis. A schematic illustration of this graph is shown in **Figure 4.** The direction of the arrows in this figure indicates the implied causal influence from RHS to LHS of equations I to VI. This graph also provides two examples of long-range transfers, namely:

a. (Fig. 4A) between conduit systems: the compression of renal vessels by the expanding pelvicalyceal conduits in the renal hilum is depicted in the sequence of transfer of pressure from the urinary fluid in the pelvis and calyces to the tissue fluid in the connective tissue linking blood vessels to the urinary tubing, subsequently compressing renal vasculature thereby reducing blood flow.
b. (Fig. 4B) along a conduit system: the flow of urine from kidney to bladder is represented over three distinct locations, namely: (i) nephron and collecting systems, (ii) ureter, and (iii) ureterovesical junction. These contiguous locations represent the entire epithelial tract connecting Bowman's capsule to bladder.

The edges of the above graph represent pairwise transfers from one location to another. To determine the route between the two locations, however, an independent computer-readable topological model of kidney structure is required

to determine if, at the level of granularity of this model, the two locations are contiguous. If the two locations share a border according to the structural model, then the route for transfer consists of the union of the two contiguous locations. If there is no contiguity between the two sites, then path-finding calculations over the structure of kidney conduits is necessary to determine path through other contiguous locations that constrain the transfer.

3) Blueprint: generalized requirements for pathophysiology pathways

So far, we have described a graph of transfers relevant to the mechanism of a single pathology. On a more general level, Point A above outlined a spectrum of transfer scenarios drawn from normal physiology as well as disease. As this spectrum of physical dynamic systems are also amenable to interpretation and representation in terms of energetic operations, the provision of a coherent computational blueprint to record these graphs is a key first step to developing a resource of physiology and pathology mechanisms shared by the Physiome and Clinical communities. The development of a Resource for Pathophysiology Mechanism (**RPM**) needs to address the following high-level requirements in its blueprint:

a. The RPM will manage:
    i. knowledge about anatomical location for the annotation of located state and rate measurements, and the routes of transfer that link these measurements;
    ii. information about experimentally or clinically determined correlations of sets of measurements that are to be represented as transfers;
    iii. documentation and provenance data about publications or peer-reviewed mathematical models from which the above correlations are derived;
b. At the core of the RPM , there must be an open and independent topological model of multiscale anatomy that provides formally-represented knowledge of different conduit systems (*e.g.* epithelial, endothelial and neural, as discussed in (21)). This model must:
    i. supply location knowledge across any scale for the annotation of state/rate measurements;
    ii. contain a reference communication map of standard transfers for the principal hydraulic systems for body fluids, such as: blood, lymph,

cerebrospinal fluid, urine, bile, chyme/chyle, pulmonary air, tissue fluid and cytosol;

iii. support the representation of population variation in the arborisation patterns of the above hydraulic and pneumatic systems;

iv. semantically link locations with community-supported reference ontologies of biological structure (*e.g.* (22)(23)(24), see also (21) for a fuller discussion of key semantic standards) to ensure interoperability with external community resources;

v. be amenable to automated inferencing that enable the calculation of transfer routes, given a pair of locations from measurements that correlate;

An implication to points iv and v above is that any external data that is already annotated with standard reference ontologies will be, from inception, compatible with this topological model. The ontology-based annotation of legacy data that is currently not annotated by any ontology will be incentivized given the added benefits of finding novel communication relationships between locations previously not achievable by inferencing directly over the standard reference ontologies.

c. Provide a formalism for the representation of transfer between anatomical locations that is able to address both the quantitative character of mathematical correlations as well as the qualitative character of located measurement defined against the reference model. One such formalism under investigation is the Bond Graph approach, which also allows for a graphical representation of a physical dynamic system (*e.g.*(25)(26)(27)(28)). A current difference is that the edges indicating transfer in Fig. 4 are unidirectional, while the Bond Graph formalism expects transfers to be bi-directional.

d. The RPM must provide shared and open tools for the two communities to:

    a. maintain and contribute knowledge to the topological model of multiscale anatomy;

    b. maintain and contribute transfer graphs representing documented correlations between located state/rate measurements;

    c. infer routes between anatomical location and visualize relevant knowledge from the topological model of multiscale anatomy and transfer graphs;

d. link to external resources such as disease terminologies, gene expression data or mathematical models in community repositories.

In practice, the technical solution may take the form of an application program interface (API) for core methods that articulate and maintain the topological model, such that calls to this API can be embedded within specialist tools serving either community.

Conclusion

The biomedical meaning of electronic health record (**EHR**) content is organized using disease terminologies such as SNOMED-CT (29)(30). One avenue of semantically bridging the biomedical meaning of Physiome models with that of EHRs is to develop a map between disease terms and variables in process models. The hydronephrosis (SNOMED-CT ID D7-14106) scenario is just one example of the definition of a standard disease entity in terms of a transfer graph of located measurements. While the manual curation of a single disease mechanism is achievable, the coherent coverage of pathology mechanisms over the entirety of the SNOMED-CT disease terminology requires both a community collaborative effort, together with the appropriate tools to effect such collaboration. The requirements collected in this work set out the blueprint of such a toolkit. The implementation of such a blueprint will provide a key means for the Clinical community to explicitly contribute and collect its knowledge about pathophysiology mechanisms to improve the integration of Physiome models in for training and healthcare decision support.
The implementation of the blueprint will be discussed in our future work.

Acknowledgements
The authors gratefully acknowledge grants from the Innovative Medicines Initiative (IMI) grants 115156-2 (DDMoRe) and 115568 (AETIONOMY), as well as European Union FP7 grant agreement 600841 (CHIC).

FIGURES

Figure 1. Three schematics for the anatomical routes relevant to the absorption, distribution, metabolism and elimination of (A) Propofol [diagram from (31)], (B) Valproic Acid (32) and (C) Docetaxel [diagram from (33)].

Figure 2. An example of a connected dependency series of six variables (rounded boxes) extracted from the Guyton 1992 model, linked via five equations (circles). Independent variables are linked to equation nodes via input arrows, dependent variables via output arrows (therefore, PRA is both an independent variable to Eq2 and a dependent variable to Eq1). Each equation may involve more than one independent variable (hence the dotted vertical arrows). The free-text definition associated with each variable symbol has been copied verbatim from the available documentation of the original model.

Figure 3. A sketch of depicting cardiovascular and neural anatomical routes between functionally-related compartments involved in conveying the interaction between **measurements A, B** and **C** as well as **model variables VRE, PRA, ANPR2** and **AAR** (shown in pink, see also Table 1). Here, compartments are represented in the shape of either boxes (*e.g.* adrenal medulla) or connecting lines (*e.g.* Vagus Nerve). The drawing of symmetrically duplicated compartments (*e.g.* right and left kidneys) was avoided in this diagram to reduce complexity of illustration. The following references were used in building this circuit: (34), (35), (36).

Figure 4. A transfer graph for the hydronephrosis pathophysiology scenario. The nodes in the graph consist of located measurements. The main types of measurement are Process rate (P) or Material state (M). The type of biophysical (*e.g.* pressure, mass) measurement is depicted in red in the top right corner of the node symbol. In the bottom right, the location of the measurement is indicated geometrically as conduit Wall (W) or Content (C), as well as in more detail in free text (red). The direction of the edges (green) indicates the location with respect to the equals sign in equations I to VI, from RHS to LHS. The subgraphs highlighted by regions A and B are discussed in the text. A mapping between the above nodes and the variables in equations I to VI is provided in **Figure 5**.

Figure 5. A mapping between the above nodes and the variables in equations I to VI.

TABLES

Dataset or Variable	Anatomical Compartment
A	The portion of blood in the lumen of the segment of internal carotid arteries containing carotid sinuses.
B	The entire pool of chromaffin cells in both adrenal medullas.
C	The entire pool of granular cells in juxtaglomerular apparatuses in both kidneys.
VRE	The portion of blood contained in the right atrium.
PRA	The portion of blood contained in the right atrium.
ANPR	The entire wall of the right atrium.
ANP	Plasma membrane of myocardial cells in both atria.
ANPC	The entire pool of blood plasma.
AAR	The entire pool of afferent arterioles in both kidneys.

Table 1. A map linking physiology data and variables (see Figs 2 & 3) to their respective anatomical compartments of these located measurements.

REFERENCES

1. Nickerson DP, Ladd D, Hussan JR, Safaei S, Suresh V, Hunter PJ, et al. Using CellML with OpenCMISS to Simulate Multi-Scale Physiology. Front Bioeng Biotechnol. 2014;2:79.

2. Keating SM, Le Novère N. Supporting SBML as a model exchange format in software applications. Methods Mol Biol Clifton NJ. 2013;1021:201–25.

3. Sirinukunwattana K, Snead D, Rajpoot N. A Stochastic Polygons Model for Glandular Structures in Colon Histology Images. IEEE Trans Med Imaging. 2015 May 15;

4. Fonseca CG, Backhaus M, Bluemke DA, Britten RD, Chung JD, Cowan BR, et al. The Cardiac Atlas Project--an imaging database for computational modeling and statistical atlases of the heart. Bioinforma Oxf Engl. 2011 Aug 15;27(16):2288–95.

5. Ulriksen G-H, Pedersen R, Wynn R, Ellingsen G. How to organize for a large-scale openEHR-based Electronic Patient Record. Stud Health Technol Inform. 2015;210:808–12.

6. Magrane M, Consortium U. UniProt Knowledgebase: a hub of integrated protein data. Database J Biol Databases Curation. 2011;2011:bar009.

7. Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, et al. The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Res. 2013 Jan 1;41(D1):D456–63.

8. Gene Ontology Consortium, Blake JA, Dolan M, Drabkin H, Hill DP, Li N, et al. Gene Ontology annotations and resources. Nucleic Acids Res. 2013 Jan;41(Database issue):D530–5.

9. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets--10 years on. Nucleic Acids Res. 2011 Jan;39(Database issue):D1005–10.

10. Lennernäs H. Modeling gastrointestinal drug absorption requires more in vivo biopharmaceutical data: experience from in vivo dissolution and permeability studies in humans. Curr Drug Metab. 2007 Oct;8(7):645–57.

11. Johnson BM, Charman WN, Porter CJH. Application of compartmental modeling to an examination of in vitro intestinal permeability data: assessing the impact of tissue uptake, P-glycoprotein, and CYP3A. Drug Metab Dispos Biol Fate Chem. 2003 Sep;31(9):1151–60.

12. Vinik AI, Ziegler D. Diabetic cardiovascular autonomic neuropathy. Circulation. 2007 Jan 23;115(3):387–97.

13. Dostalek M, Akhlaghi F, Puzanovova M. Effect of diabetes mellitus on pharmacokinetic and pharmacodynamic properties of drugs. Clin Pharmacokinet. 2012 Aug 1;51(8):481–99.

14. Yu T, Lloyd CM, Nickerson DP, Cooling MT, Miller AK, Garny A, et al. The Physiome Model Repository 2. Bioinformatics. 2011 Mar 1;27(5):743–4.

15. Tsioufis C, Dimitriadis K, Thomopoulos C, Doumas M, Papademetriou V, Stefanadis C. Renal and cardiac effects of renal sympathetic denervation and carotid baroreceptor stimulation. Curr Vasc Pharmacol. 2014 Jan;12(1):55–62.

16. Thames MD. Contribution of cardiopulmonary baroreceptors to the control of the kidney. Fed Proc. 1978 Apr;37(5):1209–13.

17. Guyton AC, Coleman TG, Granger HJ. Circulation: overall regulation. Annu Rev Physiol. 1972;34:13–46.

18. Guyton AC, Taylor AE, Granger HJ, Gibson WH. Regulation of interstitial fluid volume and pressure. Adv Exp Med Biol. 1972;33(0):111–8.

19. Omar M, Abdulwahab-Ahmed A, Chaparala H, Monga M. Does Stone Removal Help Patients with Recurrent Urinary Tract Infections? J Urol. 2015 Oct;194(4):997–1001.

20. Persky L, Storaasli JP, Austen G. Mechanism of hydronephrosis: newer investigative techniques. Postgrad Semin Am Urol Assoc North Cent Sect. 1954;67–72.

21. de Bono B, Safaei S, Grenon P, Nickerson DP, Alexander S, Helvensteijn M, et al. The Open Physiology workflow: modeling processes over physiology circuitboards of interoperable tissue units. Front Physiol. 2015;6:24.

22. Rosse C, Mejino JLV Jr. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. J Biomed Inform. 2003 Dec;36(6):478–500.

23. Bard J, Rhee SY, Ashburner M. An ontology for cell types. Genome Biol. 2005;6(2):R21.

24. Ashburner M, Lewis S. On ontologies for biologists: the Gene Ontology--untangling the web. Novartis Found Symp. 2002;247:66–80; discussion 80–3, 84–90, 244–52.

25. Zadpoor A, Arshi A, Nikooyan A. A Bond Graph Approach to the Modeling of Fluid-Solid Interaction in Cardiovascuular System's Pulsatile Flow. Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf. 2005;3:2319–22.

26. Le Rolle V, Hernandez AI, Richard PY, Buisson J, Carrault G. A bond graph model of the cardiovascular system. Acta Biotheor. 2005;53(4):295–312.

27. Chen S, Zhang S, Gong Y, Dai K, Sui M, Yu Y, et al. The role of the autonomic nervous system in hypertension: a bond graph model study. Physiol Meas. 2008 Apr;29(4):473–95.

28. Gawthrop PJ, Crampin EJ. Energy-based analysis of biochemical cycles using bond graphs. Proc Math Phys Eng Sci R Soc. 2014 Nov 8;470(2171):20140459.

29. Monsen KA, Finn RS, Fleming TE, Garner EJ, LaValla AJ, Riemer JG. Rigor in electronic health record knowledge representation: lessons learned from a SNOMED CT clinical content encoding exercise. Inform Health Soc Care. 2015 Aug 21;1–15.

30. Gøeg KR, Chen R, Højen AR, Elberg P. Content analysis of physical examination templates in electronic health records using SNOMED CT. Int J Med Inf. 2014 Oct;83(10):736–49.

31. Gill KL, Gertz M, Houston JB, Galetin A. Application of a physiologically based pharmacokinetic model to assess propofol hepatic and renal glucuronidation in isolation: utility of in vitro and in vivo data. Drug Metab Dispos Biol Fate Chem. 2013 Apr;41(4):744–53.

32. Ogungbenro K, Aarons L, CRESim & Epi-CRESim Project Groups. A physiologically based pharmacokinetic model for Valproic acid in adults and children. Eur J Pharm Sci Off J Eur Fed Pharm Sci. 2014 Oct 15;63:45–52.

33. Bradshaw-Pierce EL, Eckhardt SG, Gustafson DL. A physiologically based pharmacokinetic model of docetaxel disposition: from mouse to man. Clin Cancer Res Off J Am Assoc Cancer Res. 2007 May 1;13(9):2768–76.

34. Dampney R a. L, Coleman MJ, Fontes M a. P, Hirooka Y, Horiuchi J, Li YW, et al. Central mechanisms underlying short- and long-term regulation of the cardiovascular system. Clin Exp Pharmacol Physiol. 2002 Apr;29(4):261–8.

35. Coote JH. A role for the paraventricular nucleus of the hypothalamus in the autonomic control of heart and kidney. Exp Physiol. 2005 Mar;90(2):169–73.

36. Cano G, Card JP, Sved AF. Dual viral transneuronal tracing of central autonomic circuits involved in the innervation of the two kidneys in rat. J Comp Neurol. 2004 Apr 12;471(4):462–81.