Journal Pre-proofs

Quality by digital design to accelerate sustainable medicines development

Chantal L. Mustoe, Alice J. Turner, Stephanie J. Urwin, Ian Houson, Helen Feilden, Daniel Markl, Mohammed M. Al Qaraghuli, Magdalene W.S. Chong, Murray Robertson, Alison Nordon, Blair F. Johnston, Cameron J. Brown, John Robertson, Claire Adjiman, Hannah Batchelor, Brahim Benyahia, Massimo Bresciani, Christopher L. Burcham, Javier Cardona, Ciro Cottini, Andrew S. Dunn, David Fradet, Gavin W. Halbert, Mark Henson, Pirmin Hidber, Marianne Langston, Ye Seol Lee, Wei Li, Jérôme Mantanus, John McGinty, Bhavik Mehta, Tabbasum Naz, Sara Ottoboni, Elke Prasad, Per-Ola Quist, Gavin K. Reynolds, Chris Rielly, Martin Rowland, Walkiria Schlindwein, Sven L.M. Schroeder, Jan Sefcik, Ettore Settanni, Humera Siddique, Kenneth Smith, Rachel Smith, Jagjit Singh Srai, Alpana A. Thorat, Antony Vassileiou, Alastair J. Florence

PII: S0378-5173(25)00462-4
DOI: https://doi.org/10.1016/j.ijpharm.2025.125625
Reference: IJP 125625

To appear in: *International Journal of Pharmaceutics*

Received Date: 19 December 2024
Revised Date: 17 April 2025
Accepted Date: 18 April 2025

Please cite this article as: C.L. Mustoe, A.J. Turner, S.J. Urwin, I. Houson, H. Feilden, D. Markl, M.M. Al Qaraghuli, M.W.S. Chong, M. Robertson, A. Nordon, B.F. Johnston, C.J. Brown, J. Robertson, C. Adjiman, H. Batchelor, B. Benyahia, M. Bresciani, C.L. Burcham, J. Cardona, C. Cottini, A.S. Dunn, D. Fradet, G.W. Halbert, M. Henson, P. Hidber, M. Langston, Y.S. Lee, W. Li, J. Mantanus, J. McGinty, B. Mehta, T. Naz, S. Ottoboni, E. Prasad, P-O. Quist, G.K. Reynolds, C. Rielly, M. Rowland, W. Schlindwein, S.L.M. Schroeder, J. Sefcik, E. Settanni, H. Siddique, K. Smith, R. Smith, J.S. Srai, A.A. Thorat, A. Vassileiou, A.J. Florence, Quality by digital design to accelerate sustainable medicines development, *International Journal of Pharmaceutics* (2025), doi: https://doi.org/10.1016/j.ijpharm.2025.125625

Quality by Digital Design to Accelerate Sustainable Medicines Development

Chantal L. Mustoe[a,b], Alice J. Turner[a,b], Stephanie J. Urwin[a,b], Ian Houson[a,b], Helen Feilden[a,b], Daniel Markl[a,b], Mohammed M. Al Qaraghuli[a,b], Magdalene W. S. Chong[a,c], Murray Robertson[a,b], Alison Nordon[a,c], Blair F. Johnston[a,b], Cameron J. Brown[a,b], John Robertson[a,b], Claire Adjiman[a,d], Hannah Batchelor[a,b], Brahim Benyahia[a,e], Massimo Bresciani[a,b], Christopher L. Burcham[f], Javier Cardona[a,g], Ciro Cottini[h], Andrew S. Dunn[a,b], David Fradet[a,b], Gavin W. Halbert[a,b,i], Mark Henson[j], Pirmin Hidber[k], Marianne Langston[j], Ye Seol Lee[a,l], Wei Li[a,e], Jérôme Mantanus[m], John McGinty[a,g], Bhavik Mehta[a,b,n], Tabbasum Naz[a,b], Sara Ottoboni[a,g], Elke Prasad[a,b], Per-Ola Quist[o], Gavin K. Reynolds[p], Chris Rielly[a,e], Martin Rowland[q], Walkiria Schlindwein[r], Sven L. M. Schroeder[a,s,t], Jan Sefcik[a,g], Ettore Settanni[a,u], Humera Siddique[a,b], Kenneth Smith[a,b], Rachel Smith[a,v], Jagjit Singh Srai[a,u], Alpana A. Thorat[w], Antony Vassileiou[a,b], and Alastair J. Florence[a,b]*

[a]*CMAC, University of Strathclyde, Glasgow, G1 1RD, United Kingdom*

[b]*Strathclyde Institute of Pharmacy & Biomedical Science, University of Strathclyde, Glasgow, G4 0RE, United Kingdom*

[c]*WestCHEM, Department of Pure and Applied Chemistry and Centre for Process Analytics and Control Technology (CPACT), University of Strathclyde, Glasgow, G1 1XL, United Kingdom*

[d]*Department of Chemical Engineering, Sargent Centre for Process Systems Engineering, Imperial College London, London, SW5 2AZ, United Kingdom*

[e]*Department of Chemical Engineering, Loughborough University, Loughborough, LE11 3TU, United Kingdom*

[f]*Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, USA*

[g]*Department of Chemical & Process Engineering, University of Strathclyde, Glasgow, G1 1XJ, United Kingdom*

[h]*Chiesi Farmaceutici S.p.A. Parma, Italy*

[i]*Cancer Research UK Formulation Unit, Strathclyde Institute of Pharmacy & Biomedical Science, University of Strathclyde, Glasgow, G4 0RE, United Kingdom*

[j]*Takeda Pharmaceuticals International Co., Cambridge, MA, 02139, USA*

[k]*F. Hoffmann-La Roche AG, Basel, CH-4070, Switzerland*

[l]*Department of Chemical Engineering, University College London, London, WC1E 7JE, United Kingdom.*

[m]*UCB S.A., Brussels, Belgium*

[n]*Siemens Industry Software Limited, London, W6 7HA, United Kingdom*

[o]*Operations, Pharmaceutical Technology & Development, Sustainable Innovation & Transformational Excellence (xSITE), AstraZeneca, Södertälje, SE-151 85, Sweden*

[p]*Sustainable Innovation & Transformational Excellence (xSITE), Pharmaceutical Technology & Development, Operations, AstraZeneca, Macclesfield, SK10 2NA, United Kingdom*

[q]*Pfizer Ltd., Discovery Park House, Sandwich, Kent CT13 9ND, United Kingdom*

[r]Leicester School of Pharmacy, De Montfort University, Leicester, LE1 9BH, United Kingdom

[s]School of Chemical and Process Engineering, University of Leeds, Leeds LS2 9JT, United Kingdom

[t]Diamond Light Source, Didcot, Oxon, OX11 0DE, United Kingdom

[u]Institute for Manufacturing, Department of Engineering, University of Cambridge, Cambridge, CB3 0FS, United Kingdom

[v]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S10 2TN, United Kingdom

[w]Pfizer Worldwide Research and Development, Groton, Connecticut 06340, USA

*Corresponding author; Email: alastair.florence@strath.ac.uk

Abstract

We present a shared industry-academic perspective on the principles and opportunities for Quality by Digital Design (QbDD) as a framework to accelerate medicines development and enable regulatory innovation for new medicines approvals. This approach exploits emerging capabilities in industrial digital technologies to achieve robust control strategies assuring product quality and patient safety whilst reducing development time/costs, improving research and development efficiency, embedding sustainability into new products and processes, and promoting supply chain resilience. Key QbDD drivers include the opportunity for new scientific understanding and advanced simulation and model-driven, automated experimental approaches. QbDD accelerates the identification and exploration of more robust design spaces. Opportunities to optimise multiple objectives emerge in route selection, manufacturability and sustainability whilst assuring product quality. Challenges to QbDD adoption include siloed data and information sources across development stages, gaps in predictive capabilities, and the current extensive reliance on empirical knowledge and judgement. These challenges can be addressed via QbDD workflows; model-driven experimental design to collect and structure findable, accessible, interoperable and reusable (FAIR) data; and chemistry, manufacturing and control ontologies for shareable and reusable knowledge. Additionally, improved product, process, and performance predictive tools must be developed and exploited to provide a holistic end-to-end development approach.

1. Introduction

1.1 Background

Pharmaceutical development encompasses all the steps required to transform an active pharmaceutical ingredient (API) into a safe, effective medicine capable of being manufactured robustly and repeatably to the required quality for clinical and commercial scales. With the accelerating pace of drug discovery (Conroy, 2023), it is vital that product and process development keeps pace if we are to translate the advances in medical sciences to patient benefit as quickly, affordably and sustainably as possible. However, medicines manufacturing remains a long, complex and resource-intensive process (Schlander et al., 2021). Furthermore, approximately 90% of clinical drug development programmes may be unsuccessful due to a combination of poor clinical efficacy, issues with toxicity, weak drug-like properties, insufficient commercial requirements and/or strategic planning (Dowden and Munro, 2019; Harrison, 2016; Sun et al., 2022). As such, there is an urgent need to reduce

material wastage, energy, manpower and time during chemistry, manufacturing and control (CMC) development. It is necessary to find ways to develop new products more efficiently and robustly by exploiting the advances being made in materials, analytical, pharmaceutical and process science and engineering. In conjunction with this, it is vital to exploit advances in data science and digital technologies. Due to the range of terms presented herein, a glossary of terms is provided in Section 4, and the first instance of each term in the glossary is underlined in this manuscript. With new innovative digital approaches, it is possible to build on the well-established principles of Quality by Design (QbD) (Juran, 1992; International Council for Harmonisation., 2008).

1.2 Quality by Design

QbD is a concept introduced by Joseph M. Juran, a pioneer in quality management, who asserted that quality should be integrated into the product's design, as most quality issues stem from the initial design phase (Juran, 1992). QbD (International Council for Harmonisation., 2008) was originally proposed for use in pharmaceutical manufacturing by the U.S. Food and Drug Administration (FDA) in 2002 (Caphart et al., 2006; Department of Health and Human Services U.S. Food and Drug Administration, 2007), and was formally endorsed in 2009 (International Council for Harmonisation, 2009). QbD principles have been increasingly adopted in the pharmaceutical industry to embed quality in drug design with rigorous, science-driven approaches, and an understanding of material and processing factors and associated risks that may impact on product performance and patient safety (Azad et al., 2021; Barshikar, 2019; Davis and Schlindwein, 2018; Yu et al., 2014). It is vital to define the extent of variability in input materials that can be accommodated without impacting quality. Once the critical quality attributes (CQAs), critical material attributes (CMAs) and critical process parameters (CPPs) are identified and their interdependencies are characterised through targeted investigations, effective control strategies to manage potential risks of deviations can be developed.

QbD has five central components that have been described extensively in the literature and regulatory guidance and/or case studies (Yu et al., 2014). These can be summarised as:

1. A quality target product profile (QTPP) to pinpoint appropriate CQAs
2. Product design and understanding by performing a risk assessment to link CQAs to clinical safety and efficacy
3. Process design and understanding by defining CPPs, including a robust knowledge of scale-up and the impact of variations in CPPs and CMAs on CQAs
4. Development of a control strategy derived from the product and process understanding that ensures safety and efficacy (Lakerveld et al., 2015)
5. Process qualification to demonstrate that the controls are effective and to ensure continued improvement over time as new knowledge becomes available (Barshikar, 2019; Davis and Schlindwein, 2018)

The FDA and the International Council for Harmonisation (ICH) have published detailed guidelines on the implementation and regulatory aspects of QbD (Table 1) to help drive adoption and promote common practices for regulatory acceptability across the sector (Cook et al., 2014).

Table 1: Examples of FDA and ICH initiatives to demonstrate QbD and promote adoption and harmonisation.

Source	Initiative	Purpose	Reference(s)
U.S. Food & Drug Administration	Pharmaceutical Current Good Manufacturing Practices (CGMPs) for the 21st Century, 2002 (final report 2004)	• First mention of QbD • Promotes early adoption of technological innovation and quality systems approaches throughout pharmaceutical manufacturing. • Promotes risk-based approaches	(Department of Health and Human Services U.S Food and Drug Administration, 2004)
	Pilot program on QbD in relation to CMC, 2005	Companies invited to demonstrate QbD use in the context of CMC with a view to establishing a comprehensive quality overview of drug development and use of ICH Q8, Q9 and Q10.	(Bala et al., 2014; Department of Health and Human Services U.S. Food and Drug Administration, 2007)
	Pharmaceutical Quality for the 21st Century A Risk-Based Approach Progress Report	Improvement of the CGMP process and quality review and regulation	(Department of Health and Human Services U.S. Food and Drug Administration, 2007)
International Conference of Harmonisation	ICH Q8(R2)	Pharmaceutical Development	(International Council for Harmonisation, 2009)
	ICH Q9 (R1)	Quality Risk Management	(International Council for Harmonisation, 2005)
	ICH Q10	Pharmaceutical Quality	(International Council for Harmonisation., 2008)
	ICH Q11	Development and Manufacture of Drug Substances (Chemical Entities and Biotechnological/Biological Entities)	(International Council for Harmonisation, 2012)
	ICH Q12	Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management	(International Council for Harmonisation, 2019)
	ICH 2020 Annual Report revision of M4Q(R1)	The Common Technical Document for the Registration of Pharmaceuticals for Human Use	(ICH Secretariat with the ICH Management Committee and MedDRA Management Committee, 2021; International Council for Harmonisation, 2020)

	ICH Q13	Continuous Manufacturing of Drug Substances and Drug Products	(U.S. Department of Health and Human Services et al., 2023)
	ICH Q14	Analytical Procedure Development	(International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, 2023)

1.3 Digital Transformation of Pharmaceutical Development for CMC

With advancing capabilities in industrial digital technologies including Industry 4.0 (Arden et al., 2021; Azizi et al., 2023; Popov et al., 2022), Industry 5.0 and the underpinning engineering and physical sciences, it is increasingly possible to develop digital design methods that enable a systems-level approach to CMC development (European Commission and Directorate-General for Research and Innovation, 2022). Whilst gaps in current modelling and predictive capability remain (Hausberg et al., 2019; Kitsios and Kamariotou, 2021; Litster and Bogle, 2019), ongoing research and progress in data science and computational capabilities are addressing these issues. The potential benefits from the digital transformation of CMC include accelerated development for sustainable products and processes, risk reduction via accurate predictive computational mapping of design spaces, cost reduction via leveraging predictive tools to reduce experimental work, efficient scale-up, digital technology transfer, and cost-effective operations and patient benefit (Conroy, 2023).

Digital transformation is vital to realise the ambitions and benefits of Industry 5.0 (Akundi et al., 2022; European Commission, 2021a, 2021b, 2020; Ghobakhloo et al., 2023; Jafari et al., 2022; Maddikunta et al., 2022; Nahavandi, 2019; Rane, 2023; Valette et al., 2023; Xu et al., 2021). Industry 5.0 aims to exploit the plethora of industrial digital technologies and predictive tools such as computational and data infrastructure, digital twins (mechanistic, data driven artificial intelligence (AI)/ machine learning (ML), and hybrid models), findable, accessible, interoperable and reusable (FAIR) data principles, semantic tools, collaborative robotics, the internet of things (IoT), and immersive environments (augmented and virtual realities) (Hole et al., 2021). The goal is to ensure manufacturing is sustainable, resilient and human-centric, placing people at the heart of manufacturing with technology augmenting human creativity to create social benefit (Akundi et al., 2022). Additional drivers for innovation include the need for enhanced supply chain discernability for more robust decision making, cost-reduction, and improved efficiency via data collection and analysis across the value chain (Bermingham, 2023). Digital transformation is also being acknowledged and adopted by regulatory bodies (see ESI1) and is increasingly accepted as vital for firms to maintain their place within a competitive and dynamic global market. Furthermore, industrial digital technologies have growing potential to impact the way pharmaceutical companies design, develop and manufacture drugs (Destro and Barolo, 2022; Romañach et al., 2023; Urwin et al., 2023).

To achieve this Industry 5.0 vision, there must be a defined roadmap driven by business value, a robust technology literate workforce, and scalable, distributable technologies with secure, FAIR data systems (Wilkinson et al., 2016). Through improved exploitation of data to inform decisions, there is potential to realise environmental benefits through leveraging existing data

and reducing the reliance on resource-intensive experimental efforts (Hughes et al., 2023; Wise et al., 2019).To support the rollout and adoption of digital technologies, actions have been proposed to assist organisations in their digital transformation (Golub et al., 2023; Hauck, 2024; Whatfix, 2024):

1. Standardise and automate data collection, data control, data ownership, access management and development priorities using collaborative systems, skills and techniques, enabling better error detection and improved processes.

2. Industrialise application of AI and ML through development and adoption of standards to increase reliability of development, rollout and monitoring (Medicines & Healthcare products Regulatory Agency, 2024, 2023; Medicines & Healthcare products Regulatory Agency and Brunel University, 2022; U.S. Department of Health and Human Services et al., 2016; U.S. Food & Drug Administration et al., 2024, 2023, 2021).

3. Recruit, develop and maintain a 'tech-savvy' workforce capable of driving innovation through knowledge of digital technologies.

4. Upskill existing employees to the required level of technical expertise and provide continued development. Personalise digitalisation to address business and employee requirements.

5. Enhance digital technologies based on employee and client feedback.

6. Encourage collaboration across the sector with industry, academia and regulators working to develop standards for suitability, interpretability and credibility of AI and other digital technologies.

7. Establish a digital lifecycle approach by defining targets, identifying launch strategies for the global market and creating robust supply chains (GlaxoSmithKline plc., 2022; Golub et al., 2023; Hauck, 2024; Whatfix, 2024).

Digital tools are becoming increasingly sophisticated, user-friendly, and more accessible to researchers, manufacturers and decision-makers. This is exemplified by the recent surge in the use of generative pre-trained transformer (GPT) large language models, which have been transformative in providing non-experts with the benefits of AI tools (Eloundou et al., 2023; Mollick, 2022; Open AI, 2023). To navigate data integrity and security challenges presented by open access GPTs, pharmaceutical companies have started to implement internal solutions to incorporate AI trained on corporate and sector-specific data and improve their ways of working (Beckmann, 2023; Candelon et al., 2023). Other examples of digital tool implementation include the growing use of automation in synthetic organic chemistry, which has rapidly increased the rate of candidate throughput in discovery phases (Wang et al., 2020). In addition, computer-aided drug discovery has been used for decades to virtually screen libraries of known compounds against newly identified biological receptors to identify new possible therapeutics (Gasteiger, 2020; Medina-Franco, 2021). The addition of modern AI tools recently demonstrated efficient automated recipe planning and experimental preparation of 15 drug or drug-like substances (Coley et al., 2019; Liu et al., 2023). From an economic perspective, the use of AI tools has an estimated expected value of $15-28 billion (USD) per annum in R&D (Candelon et al., 2023). This estimation is due to an expectation that the use of AI tools will increase productivity by speeding up the process of NCE target identification, accelerating screening, and optimising formulation and product development (Adabala Viswa et al., 2024).

For CMC, ML tools also have the potential to streamline process and product development and increase overall development speed (Conroy, 2023; Nagaprasad et al., 2021). For example, to rationalise experimental solvent screening, which is expensive and time-consuming, it is only necessary to evaluate the handful of promising candidates identified by ML (Brown et al., 2018; Urwin et al., 2023). Digital-first approaches can also enable the identification of potential risks early in the development process to enable more efficient experimentation and reduce experimental load. Several studies have demonstrated the ability of digital-first approaches to support decision making, e.g., selection of manufacturing route, optimal process conditions and additional experimental conditions (Abasi et al., 2021; Agrawal et al., 2023; Bhalode et al., 2022; Boobier et al., 2020; Carou-Senra et al., 2023; Chen, 2024; Li et al., 2023; Maclean et al., 2024; Matić et al., 2023; Moreno-Benito et al., 2022; Mswahili et al., 2021; Nielsen et al., 2020; Ong et al., 2022; Pereira Diaz et al., 2023; Ravendra Singh, 2024; Sansare et al., 2021; Szilágyi et al., 2022; Urwin et al., 2023; Zaborenko et al., 2019) (Table 2).

Table 2: Examples of digital-first approaches relevant to CMC development processes.

Study Type	Technique used	Issue	Outcome and Benefits	Reference
Crystallisation	ML	Particle size quality requirement	Effective wet milling process development, reducing material wastage and experimental load, demonstrated use by Agrawal et al. (Agrawal et al., 2023)	(Urwin et al., 2023)
Hot melt extrusion	QbD and DT	Conveying, pressure build-up, and power consumption	Improved product quality so thus less wastage; highlights the need for accurate DTs and *in silico* process development	(Matić et al., 2023)
Particle processes	Hybrid ML/first principles	Particle phenomena kinetics	Successful prediction and hybrid model comparable to non-hybrid modelling in terms of accuracy, live training of the model also possible	(Nielsen et al., 2020)
Nanomedicine solubility	AI/ML Hybrid (Support Vector Regression (SVR), Multilayer Perceptron (MLP), and Least Absolute Shrinkage and Selection Operator (LASSO))	Relationship between pressure, temperature and supercritical CO2	MLP successfully predicted supercritical CO2 density, but SVR was more effective in predicting mole fraction	(Chen, 2024)

Dissolution testing	First principles/empirical approaches	Dissolution optimisation	First principles effectively provide guidance for formulation and process development, and then, in combination with data-informed empirical models, can predict the effect of material attributes (MAs) and process parameters (PPs) on dissolution for batch or real-time release	(Zaborenko et al., 2019)
Dissolution performance of direct compressed tablets	Compartmental disintegration and dissolution model (population balance principles)	Dissolution optimisation	Successfully predicted the effects of manufacturing conditions on disintegration. Utilised raw material properties as MAs and thus did not require large datasets. Suited to both DP development and live testing control strategies	(Maclean et al., 2024)
Continuous direct compression	First principles models, discrete elemental methods, flow sheet modelling, DT	Lengthy development times, material/ energy wastage, early project de-risking, improved quality	Mapping the steady state operation in the design spaces for feeding, mixing and compression. Determination of the impact of operating conditions, material and process parameters, and the dynamic response to disturbances. In turn this was used to de-risk and optimise drug product and process development while reducing the number of experiments	(Moreno-Benito et al., 2022)

Despite the advantages provided by physics-based mechanistic models, digital twins (DTs), statistical modelling, ML, and hybrid mechanistic-data driven approaches, the integration of these tools within pharmaceutical manufacturing remains bespoke. Isolated digital tools are usually only implemented on a single manufacturing step or a number of steps rather than the whole end-to-end (E2E) process from drug substance (DS) manufacturing to drug product (DP) manufacturing (Destro et al., 2021; Moreno-Benito et al., 2022; Ottoboni et al., 2022; Szilágyi et al., 2022), or for post-development optimisation (Içten et al., 2020; Liu and Benyahia, 2021) (Table 2). Developing a cohesive strategy for the digitalisation of pharmaceutical control strategies carries potential risks and requires significant investment of time and resources and working to improve regulatory certainty in the use of new methods (Ahluwalia et al., 2022). Digitalisation challenges also include risks such as cybersecurity and software issues, misinterpretation of data, and lack of training for implementation (Alguliyev et al., 2018; Axelrod, 2013; Bolbot et al., 2019; Humayed et al., 2017; Tyagi and Sreenath, 2021; Yaacoub et al., 2020).

1.4 Scope

Here, we propose a Quality by Digital Design (QbDD) framework as a holistic patient-centric approach for the integration of digital tools and data across process and product development and the manufacturing and packaging of pharmaceuticals. QbDD involves the application of extensive modelling, data driven decision support tools and generative AI-powered agents (Boiko et al., 2023; Bran et al., 2024; Chen et al., 2023; Dong et al., 2024; Mahjour et al., 2023; Ramos et al., 2024; Ruan et al., 2024; Zhou et al., 2024), to quickly, robustly and sustainably drive an efficient development procedure. This approach commences with predetermined objectives and focuses on process and product understanding and control with a strong basis in rigorous science and quality risk management (International Council for Harmonisation, 2009; Yu et al., 2014). QbDD also enables easy multi-objective optimisation by including unconventional product development objectives, for example carbon footprint or manufacturing cost reduction. Existing quality management systems in CMC include quality by testing (QbT), quality by control (QbC) and QbD. The proposed QbDD principles do not replace QbC and QbD but rather build upon their established approaches and technologies to deliver a systems-level digital transformation to CMC regulatory processes for medicines development (Figure 1). In parallel, the need for physical experimentation will decrease as digital transformation progresses towards the integration of cyber-physical systems (CPSs), with modelling and experiments (Figure 2). In addition, physical experimentation will be directed towards supporting robust outputs of the digital twin modelling. Establishing this systems-level, holistic approach will allow more efficient and effective process and product design and development to deliver the required product performance whilst also delivering improved sustainability and resilience in medicines manufacturing.

This paper is written from the perspective of small molecule, oral solid dosage form CMC process development, but the intent, approach and benefits are applicable across the full scope of medicines development.

2. Quality by Digital Design Roadmap

2.1 Towards QbDD

QbDD is enabled by the integration of digital technologies (including AI, ML and DTs) to collect and curate data, to train and apply predictive models, to enable reverse engineering and to use extended ontologies to manage and provide insights from data more effectively. These components are unified by a holistic workflow that structures development activities. QbDD uses the collated data to inform early qualitative and quantitative decisions on process selection, process design, risk assessment and control strategies. By focusing on high quality data collection, structure and connectivity as well as model development, training and validation, QbDD facilitates the identification and exploration of more robust multi-dimensional design spaces by incorporating quality, manufacturability and sustainability criteria in early-stage decision-making. Patient-centric requirements (i.e., dose, route of administration and dosing frequency) dictate the objective criteria, and the design space is minimised early in development using predictive models (Figure 2).

The exploration of a digitally augmented knowledge space in QbDD as opposed to the more sparse experimental design space in classical QbD (Figure 2) has inherent advantages in terms of efficiency. Limited exploration of the knowledge space in traditional QbD may restrict

further development. For example, in product lifecycle management, adding another presentation, such as a modified release tablet to complement an immediate release tablet, is more challenging with the reduced knowledge space of QbD. By contrast, in QbDD, the augmented knowledge space could enable screening for multiple QTPPs (including ones classically considered during product life cycle management). For example, it would be possible to screen within the same development cycle for a product presentation for acute and chronic treatment rather than considering the second treatment later in another development cycle. QbDD in general, and the exploration of the augmented knowledge space in particular, could also accelerate submission of the regulatory dossier by demonstrating to health authorities the full scope design space assessed while digital technologies populate sections of the dossier upfront.

2.2 QbDD Framework: An Integrated Predictive Toolbox

QbDD relies on establishing a holistic, systems-level framework to incorporate requirements across DS and DP development in a manner that informs agile, objective and explicit decision-making (Figure 3). This framework enables systems-level modelling and optimisation to be applied across process design, process parameter selection, molecular systems, material attributes, bulk material properties, formulated product structure-function-process interactions, stability and biopharmaceutical performance in relevant patient sub-populations (Riaz Ahmed et al., 2022). For example, predictive models can be used to design a particle formation process (i.e., crystallisation, spherical agglomeration or combined crystallisation/milling) to deliver the desired particle attributes. This targeted approach may save time and resources when compared with the traditional approach of making different samples with a range of properties and testing each to see which ones are best. Thus, models could enable efficient delivery of the desired material attributes (e.g., flowability) and assign appropriate excipients to enable effective continuous direct compression. In this way, the QbDD strategy minimises the risks and resource requirements associated with downstream processes enlarging the design space (Butters et al., 2006). Thus, the systems-level integration of QbDD allows for earlier, whole design-space refinement in place of ‘one process at a time’ development.

While QbDD will rely on the development of new digital tools, it will also utilise and build upon existing models and approaches. Examples of existing tools include, the manufacturing classification system (MCS) (Leane et al., 2015), the biopharmaceutical classification system (BCS), the developability classification system (DCS) (Amidon et al., 1995; Butler and Dressman, 2010), and process systems engineering absorption, distribution, metabolism and excretion (ADME) and pharmacokinetics/dynamics (PKPD) frameworks, e.g., gPROMS FormulatedProducts (Siemens, 2025a, 2025b), GastroPlus® (Parrott and Lavé, 2002) and SimcypTM (Certara, 2023; Jamei et al., 2009; SimulationsPlus, 2023). Additional existing models include reaction kinetics modelling (Grom et al., 2016; Wang et al., 2020), population balance models for crystallisation (Aamir, 2010; Ma and Roberts, 2019), filtration models (Ottoboni et al., 2022), wet granulation models (AlAlaween et al., 2016; Bellinghausen, 2020; Ismail et al., 2019; Jang et al., 2020) and direct compression modelling (Bekaert et al., 2022; Dai et al., 2019; Martin et al., 2021). The use of these predictive models to inform experimentation is inherently iterative. Model input and model parameters are refined, and physical experimentation feeds back into these digital tools. This partnership between digital and physical is an inherent part of the QbDD strategy and will be explored in the following sections.

Informed by industry and academia, an expanded set of tools is proposed to enable QbDD and the digital transformation of CMC with a view to reducing material and energy wastage and increasing overall sustainability across process and product development (CMAC, 2021). Specifically, three predictive modelling toolboxes are proposed covering all stages of product and process development: a biorelevant performance classification system (BPCS), an advanced manufacturing classification system (MCS+) that builds advanced simulation capability into the existing MCS, and a crystallisation classification system (CCS).

The BPCS is being developed to build on the existing BCS and DCS by connecting these classification systems to physiological- and population-based pharmacokinetics models (Amidon et al., 1995; Butler and Dressman, 2010). The BPCS will categorise APIs and formulations based on the effective range of API release achievable within a given population (Abuhassan et al., 2024, 2022a, 2022b; Prasad et al., 2022; Silva et al., 2023). In the longer term, this classification system should also aim to cover the development of dosage forms using models that self-learn from clinical outcomes and/or endpoints.

The MCS categorises drug products based on processing route and is governed by properties of the API and the needs of the formulation (Leane et al., 2024, 2018, 2015). The MCS+ is being developed to build upon the MCS by using particle and bulk property assessment to first predict outcomes for critical DP unit operations (e.g., blending, granulation and compression) which in turn enables early decision-making for DS process development (with respect to targets such as particle size and morphology) (Azad et al., 2021). Digital tools that are applicable to MCS+ relate particle properties to manufacturability targets, such as flowability, compressibility or stability, in order to guide decisions in product and process development. Predictive tools relating to excipient choice and DS interaction with excipients are also of interest for the MCS+, particularly for understanding the edges of the design space when defining an operating window that does not compromise dissolution and bioavailability.

The CCS is being developed as a predictive system of models spanning molecular properties, crystallisation thermodynamics, solvent interactions and kinetics through to particle version and form selection and physical and bulk properties. As these properties impact manufacturability and choice of formulation process, being able to predict them computationally is particularly advantageous. Thus, the CCS targets prediction of manufacturability, stability and performance parameters. Example digital tools include: estimating flow function coefficients from predicted particle size distribution (Pereira Diaz et al., 2023), optimal solvent selection to achieve desirable particle attributes (Nakapraves et al., 2022), cocrystal prediction (Devogelaer et al., 2020; Gröls and Castro-Dominguez, 2021; Loschen and Klamt, 2015), avoidance of undesirable solvate formation (Bhardwaj et al., 2015) and minimising environmental impact from solvent selection (Henderson et al., 2011). These and other models in the CCS will benefit from improvements in fundamental mechanistic understanding (Warzecha et al., 2020) as well as from access to comprehensive data sets collated from data-rich experiments exploring the response of molecules under different process conditions. To make the CCS a reality, an improved understanding of the complex interaction between solvents and process conditions and the degree to which molecular properties can inform predictions is also needed. Indeed, analysis of crystallisation kinetics through absolute rate theory signposts how molecular properties of solvents and solutes can in the future be integrated into predictive process models (Schroeder, 2024). A fully developed CCS will enable *in silico* process and particle design by guiding solvent selection, estimating

kinetics and predicting physical form, particle attributes and impurity rejection as a function of molecular properties and process conditions.

2.3 QbDD Framework: Building a Digital Workflow

The proposed QbDD digital-first approach for a given API (Figure 3) sets the overall product development objectives by identifying QTPP requirements such as dosage form, dosage, dosing frequency, delivery mechanism and pharmacokinetic parameters that ensure product safety and efficacy. Following this, prior knowledge of the API target is collated, and *in-silico* predictions are carried out to produce an initial data package of predicted and measured physical properties that may include aqueous solubility, permeability, pKa, chemical stability, anticipated impurities and target impurity profiles (*Stage 1* in Figure 3). This information is then input to the BPCS, MCS+ and CCS to explore likely outcomes from different processing routes and conditions and to target subsequent modelling and experimental efforts.

The BPCS is used to predict the target solid form, dosage, API particle attributes and formulation characteristics required to achieve safe and effective release within the target population. Coupled with information on particle and bulk properties, BPCS outputs are then fed into the MCS+ to inform both the manufacturability requirements of the system and the formulation route selection/suitability. The CCS then guides process and particle design via process, solvent and process parameter selection. Thus, QbDD essentially enables an *in silico* reverse engineering approach that first identifies targets and then determines suitable biorelevant performance, manufacturability and finally API crystallisation objectives. The resulting output from the three classification systems will be recommendations for product composition and process conditions to achieve the necessary API attributes (CCS output), drug product performance and stability (MCS+ output), and performance in the target patient population (BPCS output). Process economics, manufacturability and sustainability targets are then determined as well as known constraints (*Stage 2* in Figure 3). Process options and models relevant to these targets are also identified (*Stages 3* and *4*).

Following the first four *in silico* stages in the workflow, the first experimental call outs occur in *Step 5* to enable model calibration and refinement. During development, predictive tools may require some level of experimental input and can be trained and calibrated with model-driven callouts for automated, data-rich, small-scale (<2 g or <10 mL), materials-sparing experiments. Model-driven experimental platforms (Christensen et al., 2021; Rogers et al., 2020), identify operational constraints and allow faster option assessment, process verification, process validation and data feedback to inform the overall model. Some phenomena are currently not well covered by predictive simulation tools, e.g., fouling, nucleation or mechanical properties of bulk materials, and will require further investigation and model development.

After model calibration and refinement, conceptual process and product options are investigated via model driven development (*Stage 6).* Process model validation in this stage may require physical experimentation to assess discrepancies such as those caused by non-modelled phenomena. *Stage 6* determines which process options will be taken forwards, and this is followed by an initial quality risk assessment using process models coupled with practical constraints (*Stage 7*). When combined with sensitivity analysis and holistic design

space process optimisation against the process objectives whilst ensuring quality, the risk assessment allows for the initial definition of a control strategy (*Stages 8* and *9*). The application of stages 1-9 (Figure 3) ensures that the QTPP is met and delivers the required product performance in a science-led, digital-tools-enhanced, materials-sparing and efficient manner. In *Stage 10,* the process is operated to produce material. The material and associated process analytical technology (PAT) data are analysed, and results used to determine if the QTPP is met and the model is credible. Product performance can then be assessed, relevant data is used for further model improvement, and the product lifecycle is managed with ongoing improvements facilitated by the digitally augmented knowledge space (*Stage 11* and *12).*

Subsequent iterations of the MCS+ and CCS and potentially BPCS may occur in both early and late-stage process development as further data become available. Examples of this are: (a) should API particles not be conducive to direct compression and instead result in poor blend dispersion, poor drug loading or lack of tablet uniformity, alternative routes such as roller compaction, wet granulation or solid dispersion methods must be considered; (b) should issues with end-product properties be identified during product analysis, this information is fed back to the classification systems to alter the formulation; (c) should a formulation fail to release the API in the expected manner, the excipient choice, formulation process parameters, formulation processing route or even the DS particle production route may be reconsidered. As new and improved models become available, an increasing reliance and trust in digital design methodologies should continually reduce the reliance on experimental data (as indicated by the decrease in physical and increase in virtual experiments in Figure 2), and in turn reduce any need to revisit the classification systems in subsequent stages of development.

Other feedback loops may occur throughout the QbDD workflows. For example, if a suitable operating space (i.e., one that assures quality, manufacturability and sustainability objectives within the process and operating constraints) cannot be found following model validation, it may be necessary to return to Stage 2 to assess if any objectives can be relaxed before proceeding with Stage 8. Likewise, if scale up verification and validation experiments do not meet predicted outcomes e.g., due to non-modelled phenomena, it may be necessary to revisit the models, the model parameters (*Stages 4 and 5*) or redesign the experimental set up (*Step 6*) before proceeding.

Making decisions on a quantitative basis early in development using this approach can have significant benefits for the cost, time and ultimate sustainability of medicines development. To truly realise fully predictive capabilities for early decision-making, further development and validation of predictive tools and multi-scale, multi-physics model frameworks is required (Figure 3) (Liu et al., 2021; Niederer et al., 2021; Onaji et al., 2022). However, ongoing development of increasingly capable mechanistic (Chaudhury et al., 2014; Ottoboni et al., 2022), AI (Chakravarty et al., 2021; Coley et al., 2019; Madarász et al., 2023) and hybrid models (Abouzied et al., 2023; Bhalode et al., 2022; Chen, 2024; Gaddem et al., 2024; Nielsen et al., 2020; Tsopanoglou and Jiménez del Val, 2021), across different scientific endeavours highlights both the accelerating progress and the huge potential achievable.

2.4 QbDD Elements

2.4.1 Cyber-Physical Systems

QbDD will benefit from the growing capabilities of cyber-physical systems (CPSs) that interlink computational technologies with physical processes, as part of cyber-physical research infrastructures (CPRIs), to analyse, monitor and control their functionality in a consistent, robust, safe, efficient and concurrent manner (Alguliyev et al., 2018; Baheti and Gill, 2011; Lee, 2006; Marwedel, 2021; Robotics Growth Partnership, 2022; Sanislav and Miclea, 2012). CPSs, as a concept, are now considered vital in terms of innovation across Europe and the United States (Alguliyev et al., 2018; Robotics Growth Partnership, 2022), as reflected in the drive towards Industry 4.0. With the Industry 5.0 focus on placing humans at the heart of manufacturing, to be effective in realising the envisioned transformation, CPSs should function as a collaborative synergistic human-machine interface providing well-trained, technology-savvy workers with ready access to data and models across DS and DP activities (European Commission, 2021b, 2021a, 2020; Jafari et al., 2022; Nahavandi, 2019).

QbDD could also benefit significantly from the accelerating application of self-driving laboratories (SDLs) utilising CPSs driven by data, models, optimisation approaches and AI. SDLs have demonstrated increased data acquisition (e.g., an increase of sample throughput of 50-100 times compared to human testing alone) (Berkeley Lab, 2024; Biron, 2023), novel material discovery (Burger et al., 2020; Yang and Tomoshige, 2024), real-time monitoring and analysis (Harmon, 2023), and even multicontinental DT and CPS co-ordination (Bai et al., 2024). Furthermore, multiple SDLs have demonstrated real-time recording, storage and interaction with data whilst also using human input and literature data (Acceleration Consortium and University of Toronto, 2024; Aspuru-Guzik Group, 2024; Bai et al., 2024; Harmon, 2023; Intrepid, 2024). QbDD has not yet been implemented in existing pharmaceutical SDLs in a full E2E capacity (Acceleration Consortium and University of Toronto, 2024; Berkeley Lab, 2024; Intrepid, 2024). To this end, CMAC is actively developing QbDD-enabling SDLs or DataFactories™. These DataFactories include SDLs capable of collecting targeted experimental data for APIs, excipients, and products under a wide range of conditions. They do so by exploiting automated dosing or sample handling, mobile robotics, small-scale experiments with integrated sensing/analytics/imaging for information extraction and global optimisation approaches for self-learning experimental planning to meet the process objectives. These SDLs focus on model-driven data generation via repeatable experimentation with data structured in FAIR formats (discussed further in Section 2.4.3) and currently target crystallisation screening and scale-up (Pickles et al., 2024, 2022a, 2022b), amorphous materials, direct compression (Abbas et al., 2025), stability and dissolution testing (CMAC, 2021).

2.4.2 Data Systems and Architectures

The data sources and requirements across a QbDD framework are numerous (e.g., Figure 4), and an underlying data structure is necessary to facilitate the integration, collation, management and application of data between and across these platforms. Data systems and architectures provide a standardised structure for data collection, processing, organisation, security and storage (University of York, 2024). Several local and enterprise-level data frameworks have been developed including highly ordered data warehouses and more unstructured data lakes and, more recently, data meshes and data fabrics (IBM, 2024) (please see Section 4 for full definitions). Due to issues such as lack of flexibility and lack of quality (IBM, 2024), data fabrics and data meshes may be utilised in preference to data warehouses

and data lakes (Dibley, 2022; Garani et al., 2019; Garcia et al., 2008; Hlupić et al., 2022; IBM, 2024; Nambiar & Mundra, 2022; Thantilage et al., 2023). Data fabrics and data meshes enable data to be efficiently managed and made accessible to a range of human users, applications and other systems further down the supply chain, either in a decentralised (mesh) or centralised (fabric) form (Blohm et al., 2024; Hechler et al., 2023). For the implementation of QbDD, the centralised management and accessibility of a data fabric makes it the preferred data architecture for a holistic, human-centric QbDD framework. Data fabrics enable full integration of different data sources and pipelines across different locations, allowing the collation and curation of all data and metadata to streamline access and drive modelling approaches (IBM, 2024). Important elements for a FAIR QbDD data fabric include; attributable, legible, contemporaneous, original accurate, complete, consistent, enduring and available (ALCOA+ principles) data; cybersecurity; ontologies; extract-transform-load (ETL) or extract-load-transform (ELT) tools to ensure all data can be correctly tagged, aggregated and served up to queries or data-driven services including dashboards; and AI/ML (Bartley, 2024; Durá et al., 2022; Kavasidis et al., 2023; Samson, 2021; Seenivasan Mphasis and Seenivasan, 2023; Sembiring and Novagusda, 2024; Singhal and Aggarwal, 2022).

Investment in establishing standardised data architectures (Veeva Systems Inc., 2025; Walsche, 2021) and systems will be valuable in developing and deploying modelling and data-driven approaches. Effective mechanistic digital twins and ML-based or hybrid predictive models with known uncertainties are facilitated by structured training data in which metadata describing those data are adequately captured. For example, a powder flow prediction model trained on data collected on one type of instrument with one methodology has a lower accuracy when predicting the outcomes for the same physical property measured using another instrument that uses a different measurement methodology (Pereira Diaz et al., 2023). Capturing the metadata, such as differences in methodology (e.g., equipment scales and conditions) (Wang et al., 2021), is therefore essential to improve predictive model performance and ensure interoperability and repeatability. As recently demonstrated for high shear wet granulation (Wang et al., 2021), material data fusion and multivariate modelling can also speed up process development by connecting several different data sets and reducing the volume of experimentation required. FAIR data collected from multiple sources, enriched with metadata and in machine-ready format is then appropriate for advanced analytics, i.e., ML and AI.

Recently, there have been initiatives to embrace FAIR data principles and replacing traditional data tables in databases with detailed knowledge graphs (Strömert et al., 2022; Voigt and Kalidindi, 2021; Wulf et al., 2021). A knowledge graph, consisting of an ontology and appropriate data, can be used to capture and represent knowledge and relationships between data entities in the domain, enabling semantic tool development to improve data access and usability. In general, ontologies describe classes of objects, entities or concepts and the relationships between them (Lomax, 2022). An ontology forms a machine-readable knowledge model that supports FAIR data generation through connecting data and meta data intuitively and aids the discovery of deviations, thereby decreasing errors and enabling quantification of uncertainty and confidence in data (Francisco and Remolona, 2018; Lomax, 2022; Strömert et al., 2022; Venkatasubramanian et al., 2006; Viswanath et al., 2022). Developing a complete QbDD ontology for CMC in pharmaceutical manufacturing will require significant time, resources and maintenance. To alleviate these challenges, existing ontologies must be leveraged, such as those developed for pharmaceutical engineering (covering material properties, molecular structure, experiments, reactions, phases and operations (Hailemariam and Venkatasubramanian, 2010), secondary process training (Chalortham et al., 2013; Oyebola and Opeoluwa, 2015) and the Chemical Entities of Biological Interest's (ChEBI) ontology of molecular entities developed by European Bioinformatics Institute (EBI) (Chemical

Entities of Biological Interest (ChEBI), 2024). This endeavour will also benefit from groups and organisations working together to form the basis of a standard knowledge model for the domain, driving adoption of the resulting ontology and maintaining it to ensure longer-term impact.

2.4.3 QbDD Framework: The Underlying Data Fabric

Establishing a CMC data fabric can facilitate the integration and transfer of data between the predictive models, digital twins, DataFactories (Pickles et al., 2024) and other targeted experimental data sources used to identify material attributes, process parameters and associated quality attributes. Critically, an underlying data fabric will also rely on the accessibility of data generated by multiple analytical instruments. Instrument compatibility with standardised communication protocols, such as Standardisation in Laboratory Automation (SiLA) (SiLA, 2018), and Robotic Operating Systems (ROS) (Open Source Robotics Foundation, 2024), will be more and more integral to realising this potential. Developing CMC ontologies and connected knowledge graphs to enable FAIR data will also address the industry-wide challenges of connecting data silos and data interoperability. This connectivity will benefit complex data sets spanning all stages of development for whole-process design and optimisation, sustainability metrics and intellectual property management. Data in QbDD should not be isolated by a geographical research site or operational team; instead, it must be transparent to all institution members and, potentially, regulators through federated systems (Ahluwalia et al., 2022).

The data fabric will also connect predictive models with the self-driving DataFactories required to generate data to train, calibrate and evaluate model predictions. Model-driven automated experimental frameworks investigate gaps in existing knowledge using AI and model-driven decision-making (Gregoire et al., 2023) and can develop suites of data that provide value to future campaigns and contribute to the development of improved material and process understanding (Abouzied et al., 2023; Boobier et al., 2020; Carou-Senra et al., 2023; Chen, 2024; Mswahili et al., 2021; Nielsen et al., 2020; Ottoboni et al., 2021; Patel and Shah, 2022; Pereira Diaz et al., 2023; Vassileiou et al., 2023). The ambition of such digital design efforts are that reliable, credible predictive models generated by the data fabric support the development of process-mirroring digital twins that can dynamically replicate processes and allow for analysis, control and optimisation of critical attributes in real time.

2.4.4. Interconnectivity of QbDD Elements

Figure 5 demonstrates a potentially fully integrated QbDD platform in which four CPRI platforms are brought together by the QbDD data fabric:

- The *Skills* platform of CPRI comprises multi-skilled Industry 5.0-ready researchers working in a diverse and inclusive interdisciplinary laboratory environment for innovation of medicines manufacturing.

- The *Measure* platform consists of multiple types of CPSs (A,B, and C in Figure 5). This platform is a human-centric integrated CPRI that builds on individual CPS for individual unit operations (A in Figure 5) to integrate experiments across multiple process Stages (B in Figure 5) and ultimately E2E use (C in Figure 5) with data and metadata feeding into the research data fabric (D in Figure 5).

- The *Model* platform of the CPRI is a systems-level manufacturing knowledge model for CMC where predictive models of single rate processes (E in Figure 5) and drug substance/drug product models (F in Figure 5) can be integrated into an E2E model framework (G in Figure 5) as a system-level digital twin and utilised to inform knowledge management (H in Figure 5).

- The CPRI *Make* platform is an integrated, scaled-down, material-sparing E2E customisable manufacturing research test bed to validate the quality, sustainability and resilience of adaptive processes and control strategies. In this platform, initial processes (I in Figure 5) are developed to create integrated small scale, flexible, modular continuous processing platforms (J in Figure 5), which in turn can be combined to generate E2E integrated flexible, modular continuous processing platforms (K in Figure 5) and multi-route E2E integrated flexible, modular continuous processing platforms (L in Figure 5).

2.5 Digital-First Workflows

The integration of the *Measure*, *Model* and *Make* framework (Figure 5) requires guided decision-making and handovers through digital workflows. Workflows provide a structured approach to process and product development, and the transparency provided in key decision-making reduces process risk and uncertainty (Agrawal et al., 2023). Reasoning and support for each decision are clear and accessible to different disciplines, geographies and phases of the development process, which will only benefit regulatory processes. Benefits have been reported in the development and application of workflow methodologies to improve the development of different stages of pharmaceutical development (Brown et al., 2018; Hatcher et al., 2020; Ottoboni et al., 2021; Pickles et al., 2024; Urwin et al., 2020). Additionally, in other groups (Agrawal et al., 2023; Hu et al., 2024; Lorenz et al., 2021; Sperry et al., 2021), with typical advantages being reduced development time and resource requirements.

The overarching QbDD workflow (Figure 3) drives the digital-first strategy, with an ultimate goal to exploit predictive models to rapidly identify the optimum materials, equipment and process conditions under which QTPPs can be achieved. This workflow with associated sub-workflows connect key decision points for each selected process stage with predictive models and data derived from model-driven experimental design to interrogate reaction mechanisms, parameterise models, quantify uncertainty and optimise design solutions. The QbDD data fabric ensures workflows can callout to required data across development process operations. Via DataFactories, workflows can also trigger the generation of targeted, reproducible data to drive models applied at key development stages (Pickles et al., 2024, 2022b). Workflows also integrate the *Skills* platform (Figure 5) into the QbDD framework up/reskilling users through guided decision-making and accessibility and interpretability of the data and associated models. For example, dashboards can summarise experimental progress, data trends, model performance and other useful metrics for each QbDD stage. Templates must be also integrated to facilitate FAIR data capture for sample data, operation-specific data and metadata.

2.6 Validation and Maturity of Models and Digital Technologies

Many digital platforms may follow the same basic installation, operation and performance qualifications (IQ, OQ and PQ) that standard equipment follows. Overall IQ, OQ and PQ can

ensure any models or other digital technologies, including automation, robotics and control systems, fit the design specifications. Additionally, these models and digital technologies may require monitoring and controlling using appropriate calibration and validation to achieve reliable and robust system performance (Creaner and Fitzgerald, 2024). Similarly, the cyberphysical infrastructure underpinning QbDD may require the use of verification, validation and uncertainty quantification (VVUQ) processes vital to the model life cycle. Computer Systems Validation (CSV) ensures models function as intended, consistently and accurately (The FDA Group, 2023). Model verification establishes if the model fits its mathematical description (The American Society of Mechanical Engineers, 2024). Model verification should occur throughout the product life cycle as defined by the FDA in their Appendix to Q8, Q9 and Q10, ensuring the model fulfils its acceptance criteria (U.S. Department of Health and Human Services et al., 2012). Scientific model validation establishes model accuracy relative to experimental set up (Ahmed et al., 2012; The American Society of Mechanical Engineers, 2024), as demonstrated by a number of studies (Barrasso et al., 2015; Chen et al., 2023; Kodam et al., 2012; Moreno-Benito et al., 2022; Ranjan Yadav et al., 2022; Sadeghi et al., 2022; Unnikrishnan et al., 2021). Uncertainty quantification captures the effect of variations in modelled and experimental process parameters on the output and key performance indicators (The American Society of Mechanical Engineers, 2024). Workflows and data flows can similarly be validated by assessing if workflow output matches a defined, measurable objective in a transparent, repeatable way. Workflow and data flow validation is ongoing with every iteration of the workflow and with ongoing assessment of data provenance and representation, for instance. For QbDD methods and approaches, credibility aims can be identified as part of the risk assessments. Then, a suitable verification and validation strategy can be designed and implemented to establish the overall credibility of the model, and experimental data then can be used to inform and alter the model as required (Ahmed et al., 2012; The American Society of Mechanical Engineers, 2024).

3. Outlook and Recommendations

3.1 Outlook

QbDD facilitates the digital transformation of CMC processes for medicines product and process design and manufacturing by establishing a holistic cyber-physical framework. This framework can be realised using inclusive, digitally-encoded CMC workflows that guide all development stages and objectives. The implementation of systematic workflows has been shown to improve the overall efficacy of automation and predictive models (Coley et al., 2019; Içten et al., 2020; Ottoboni et al., 2022). Furthermore, automated workflows have the potential to overcome complexity barriers, and rapidly equip experimentally trained scientists with access to digital tools for efficient process development (Golub et al., 2023; Hauck, 2024; U.S. Food & Drug Administration, 2021; Whatfix, 2024). Development of these workflows will require collaborative efforts between academia, industrial partners and regulatory stakeholders to drive standardisation, acceptance and adoption across the range of materials, processes and unit operations relevant to all dosage form development.

Further investment in the development of data infrastructure is required to drive the digital transformation of CMC. This development will enable the ability to acquire, curate, store and analyse data and/or metadata generated on-demand from SDLs, simulations and other experimental sources. In turn, this develops process understanding and our ability to train and develop models that evaluate design space, manufacturability and sustainability. SDLs and DataFactories are already being developed by a range of initiatives for pharma and non-pharma applications (Acceleration Consortium and University of Toronto, 2024; Berkeley Lab,

2024; Biron, 2023; Intrepid, 2024; Pickles et al., 2024, 2022b, 2022a; University of Liverpool, 2024), demonstrating the ability to enhance efficiency of R&D. A CMC ontology and associated knowledge graph can both ensure data collected is FAIR, maximising opportunities to extract knowledge and value and will facilitate model development and process scale-up. Standardisation of data collection, data structuring and experimentation via SDLs may also provide significant benefits through regulatory process innovation, for example simplifying assessment and review of data associated with a new product (Berkeley Lab, 2024; Biron, 2023).

Although this perspective paper focuses on the potential for QbDD across CMC product and process design, the principles can be readily extended to drug discovery, synthesis, scale-up, manufacture and life cycle management. Drug discovery and synthesis prediction, AI, digital twins, CPSs and ML can inform the CCS, MCS+ and BPCS (An and Cockrell, 2022; Blanco-González et al., 2023; Bordukova et al., 2024; Coley et al., 2019, 2018; Cumming et al., 2013; Dara et al., 2022; David et al., 2020; Elbadawi et al., 2021; Gregoire et al., 2023; Grom et al., 2016; Jayatunga et al., 2022; Jiménez-Luna et al., 2021; Jing et al., 2018; Lavecchia, 2015; Lo et al., 2018; Medina-Franco, 2021; Patel et al., 2020; Patel and Shah, 2022; Subramanian, 2020; Vamathevan et al., 2019; Wu et al., 2023; Yang et al., 2019; Zhang et al., 2017). For example, the resultant isolated compound and its associated impurities can inform crystallisation requirements, processability and overall product quality (An and Cockrell, 2022; Blanco-González et al., 2023; Bordukova et al., 2024; Coley et al., 2019, 2018; Cumming et al., 2013; Dara et al., 2022; David et al., 2020; Elbadawi et al., 2021; Gregoire et al., 2023; Grom et al., 2016; Jayatunga et al., 2022; Jiménez-Luna et al., 2021; Jing et al., 2018; Lavecchia, 2015; Lo et al., 2018; Medina-Franco, 2021; Patel et al., 2020; Patel and Shah, 2022; Subramanian, 2020; Vamathevan et al., 2019; Wu et al., 2023; Yang et al., 2019; Zhang et al., 2017). Additionally, QbDD has potential in the packaging space. In line with MCS+, CCS and BPCS, challenges such as moisture absorption, hydrolysis, material interactions, friability, light sensitivity, counterfeit products and poor patient compliance inform packaging selection (Allain et al., 2016; Bahaghighat et al., 2019; Chen and Li, 2003; Cramer, 1998; Feenstra et al., 2014; Naveršnik and Bohanec, 2008; Remmelgas, 2017; Waterman and MacDonald, 2010; Zhao et al., 2022). Thus, predictive models (Crews et al., 2018; Feenstra et al., 2012; Naveršnik and Bohanec, 2008; Remmelgas, 2017), real-time modelling (Vijayakumar et al., 2024), digital twins (Schrimpf, 2022), ML (Deshpande, 2023; Jones, 2024), and AI (Brownett-Gale, 2024; Tubettificio Perfektup, 2023), have been implemented previously. Furthermore, a similar digital-first approach has been previously suggested in RNA vaccine research, albeit not in an end-to-end capacity, demonstrating QbDD's applicability to non-solid dosage forms (Nair et al., 2024). In the broader supply chain context, digital twin supply chains (DTSCs) are also emerging (Srai et al., 2024, 2020, 2019). DTSCs have been explored in relation to synchronisation of pharmaceutical production and logistics operations (Guo et al., 2024), simulation-based capacity planning (Santos et al., 2020), inventory optimisation (Marmolejo-Saucedo, 2020), and data standardisation and integration along the value chain (Werner et al., 2021). As seen in these extensive examples, QbDD's workflow-guided integration of digital tools has potential impact beyond the scope of this publication and with continuing improvements in predictive tool availability will gain increasing momentum in the coming years.

QbDD will require ongoing investment of time and resources to develop and integrate the advanced manufacturing and digital technologies to establish mature cyber-physical infrastructure for modern CMC development. Whilst digital technologies such as AI are having impact in drug discovery (Adabala Viswa et al., 2024; Chakravarty et al., 2021), organisations also require a clear business case showing the expected return on investment and improved outcomes in CMC development and manufacturing processes. By reducing materials,

instrument time, staff time and potential exposure to more potent materials, the use of AI, ML-driven, physical and hybrid predictive models and DTs could improve overall efficiency, safety and sustainability of CMC development efforts. There is also a need to build trust and confidence in models and digital technologies and ensure data integrity and the cybersecurity of data transfer and data systems. QbDD also increases the number of product and process prototypes investigated (mostly in the virtual space), which, in turn, increases the likelihood of identifying and developing products that maximise patient benefits. International regulators including the Medicines & Healthcare Products Regulatory Agency, the European Medicines Authority and the U.S. Food and Drug Administration are adopting and encouraging digital transformation in pharmaceutical manufacturing (European Medical Agency, 2020; European Medicines Agency, 2024, 2017, 2016; European Medicines Agency and Heads of Medicines Agencies, 2020; Pauli and Williams, 2018; Riaz Ahmed et al., 2022; U.S. Food & Drug Administration, 2021, 2019; Yu et al., 2019) (see ESI 1 for further details). This guidance provides increasing momentum behind the adoption of novel, digital technologies, infrastructure and innovative ways of deploying them.

In conclusion, QbDD facilitates the efficient and robust development of processes that deliver drug products which meet quality, manufacturability, process sustainability, regulatory and business targets with enhanced understanding of these processes. The enabling technologies for QbDD are cost-effective, available and proven and combine FAIR data principles, a QbDD data fabric, predictive models and material-sparing, highly automated experimentation. Through these, QbDD provides the science-driven rationale, data transparency and traceability, more robust design space, real-time process improvements and a formalised decision process to enable organisations to enhance R&D productivity and provide regulatory confidence and assurance. Ultimately, this allows for the realisation of Industry 5.0 principles in pharmaceutical process and product development that can help to sustainably translate new medical science into new medicines to improve the lives of patients.

3.2 Recommendations for Realising QbDD

Table 3: Recommendations for Realising QbDD. The example timeline assumes ongoing application of QbD within an organisation and will differ from one organisation to another.

Recommendation	Example Timeline for Related Activities			Associated Benefits
	Near term (1-2 years)	**Medium term (3-5 years)**	**Long term (5+ years)**	
Introduction of advanced computational resources	• Identify and develop relevant ML, mechanistic and hybrid models • Verify and validate models	• Assess model credibility and regulatory readiness level (guidelines to be determined through regulatory engagement, see near-term goals) • Continue model development and validation • Implementation of models according to credibility assessment and technology readiness level	• Continue model development, validation and assessment • Continue model implementation according to credibility and lifecycle requirements	• Promotes and enables digital first approaches such as workflow guided DTs, ML, predictive models and hybrid approaches • Enhances efficiency and sustainability of CMC by early design space refinement via virtual assessment delivering approved medicines to patients faster • Reduces risk of unforeseen late-stage design space re-assessment and expansion • Reduced material wastage, energy and workforce time • Mitigates environmental impact (promoting sustainability) and reduces risk for researchers with virtual assessment of less-desirable operating conditions • Standardised approaches to evaluate model suitability and credibility for different contexts of use.
Introduction of new data technologies including: • ontologies • FAIR data approaches	• Map data interdependencies and meta-data • Identify data integration approach (internal development vs. external software solutions)	• Implement and test ontology solutions • Test and measure data FAIRness	• Expand ontology as needed • Continue to periodically test data FAIRness	• Provides additional transparency, data integrity and traceability in process and product development • Facilitates seamless data connectivity • Provide standardised framework to assess FAIR data approaches and benefits • Enables regulators to more easily assess the validity of the processes and products developed

Develop and use holistic, consistent and standardised digital workflows to guide QbDD	• Develop sub workflows for the high-level workflow presented in this work with institution-specific pathways	• Revise workflows to integrate new, credible digital tools with the goal of continued reduction of physical experimentation	• Continue workflow revision to minimise physical experimentation by further incorporation of new or better digital tools	• Establishes a comprehensive platform for QbDD. • Combines a patient-focused QTPP with route selection, modelled CPPs, minimal small-scale experimentation and rapid digitally enabled technology transfer and scale-up. • Drives and integrates digital and experimental activities. • Captures all key CMC objectives, constraints, decision points and data requirements to drive and coordinate digital and experimental activities • Rapidly equips experimentally trained scientists with digital tools for efficient process development • Accelerates the adoption of QbDD principles across pharmaceutical development and manufacturing organisations • Increases the R&D productivity and efficiency in terms of both cost and sustainability • Standardises interfaces between workflows, data, predictive data, predictive models and simulations • Delivers a more robust design space with known uncertainties and sensitivities that can be used to evaluate risks and define the control strategy
Introduction of new technologies such as self-driving labs (SDLs) or DataFactories (DFs) and integrate these platforms using workflows	• Identify areas where SDLs and DFs are needed (i.e. areas of interest for ML or hybrid model development) • Identify, procure and integrate instruments and software for FAIR data generation • Develop plan for data integration into wider ontology	• Integrate data collection with wider ontology • Continue instrument integration and platform development • Collect data and ensure continued data FAIRness	• Assess new available technologies for incorporation into existing DFs/SDLs or development of new DFs/SNLs • Assess continued viability of existing DFs/SNLs - if they have they achieved their data collection goals, can they be repurposed or can individual components be repurposed	• Enables implementation of digital-first approaches • Standardises approaches to FAIR data generation and reporting. • Allows for efficient, repeatable, data-rich and model-driven experimental data generation • Allows for targeted generation of data to train, parameterise and calibrate models and validate predictions • Significantly accelerated experimentation rate, lower amounts of material and automation reduce exposure risk for staff
Engage regulatory stakeholders (e.g. the UK-based digital CMC Centre of Excellence in	• Develop national and international guidelines to assess model credibility and	• Establish and agree upon international guidelines with multiple international regulatory bodies	• Publish harmonised international guidelines • Periodically assess guidelines with the development of new	• Encourages adoption across the sector • Provides a common language and format for regulators, industry and academia • Drives standardisation

Regulatory Science (CERSI))	regulatory readiness level • Provide regulatory-relevant training for stakeholders in model credibility and regulatory readiness level assessment	• Provide training for model developers and industry researchers on testing, implementation and integration of digital tools in industry according to agreed guidelines	and existing technologies • Continue to provide relevant trainings	
Training and upskilling existing and future workforce (this work lies within the remit of current Centres for Doctoral Training (CDTs) amongst other training programs)	• Identify, apply for and support opportunities for multi-disciplinary training with QbDD focus for current and future workforce (e.g. CDTs and training development programs and platforms). • Develop QbDD-relevant trainings such as: integration and use of FAIR data and data structures; model and SDL development and/or implementation; model credibility and regulatory readiness	• Implement, facilitate, and encourage QbDD-relevant trainings/training programs for current and future workforce • Develop training for new technologies as required	• Continue QbDD-relevant training for current and future workforce • Continue to develop training for new technologies as required	• Builds on existing knowledge and enables continued improvement of existing processes • Provides additional tech-savvy workforce
Assure all areas of development, from digital twins and predictive models to CPS	Ensure all models and platforms: • Have associated risk assessments with verification and validation control strategies • Follow IQ, OQ and PQ guidelines • Are verified throughout the product life cycle in	• Continue to assure all areas of QbDD development with actions listed under near-term goals	• Continue to assure all areas of QbDD development with actions listed under near-term goals	• Fulfils the requirements of VVUQ • Establishes the overall credibility of the models, CPS and more generally of QbDD • Enhances overall efficiency and effectiveness of medicines manufacturing whilst assuring product quality

	accordance with the ICH requirements • Are validated relative to the experimental set up • Are monitored for variations in modelled and experimental outcomes • Are fully integrated with processes and data capture			
Combine a patient-focussed QTPP, route selection, modelled CPPs, small-scale experimentation and digitally enabled technology transfer	• Collaboratively develop BPCS, CCS and MCS+ classification systems • Integrate classification systems early in workflows • Use multi-objective optimisation for sustainability and process needs whilst assuring quality and safety	• Continue and finalise development of classification systems • Continue development of mechanistic, ML and hybrid models to inform classification systems • Reduce physical experimentation as classification system maturity improves	• Continue development of mechanistic, ML and hybrid models to inform classification systems • Continue to reduce physical experimentation as models informing classification systems improve	•Allows delivery of a more robust design space with known uncertainties and sensitivities •Allows effective evaluation of risks and establishes a robust control strategy •Enables more sustainable and cheaper medicines to get to the patient faster
Incorporate PAT where required to assure quality requirements and inform model development	• Prioritise PAT first according to quality assurance and secondly to inform model development • Assess which PAT data can be used to inform model development • Develop strategy to incorporate PAT data into QbDD ontology and data structures	• Continue to prioritise PAT first according to quality assurance and secondly to inform model development • Incorporate PAT data into QbDD ontology and data structures	• Continue to prioritise PAT first according to quality assurance and secondly to inform model development	•Maintains quality assurance •Continued supply of data to the QbDD data fabric to inform models and improve overall processes in real time

4. Glossary

In this section, we provide a list of definitions and abbreviations to remove ambiguity for the terms used throughout this paper.

Active pharmaceutical ingredient (API). Any component that provides pharmacological activity or other direct effect in the diagnosis, cure, mitigation, treatment, or prevention of disease, or to affect the structure or any function of the body of man or animals.

Artificial intelligence (AI). A system which carries out computer and machine-driven problem solving, in a way that mimics human intelligence. The FDA defines AI as a machine-based system capable of providing predictions, suggestions or decisions impacting actual or simulated environments and uses information from human and machine sources to analyse these environments to generate appropriate models to exert an appropriate response (U.S. Food & Drugs Administration, 2024).

Attributable, legible, contemporaneous, original accurate, complete, consistent, enduring and available (ALCOA+) data. Defined by the FDA as the guiding principles for data integrity as laid out under CGMP, where all data must be traceable, decipherable and gathered and recorded within the appropriate time scale (concurrently if possible). Additionally, the first version must be kept, and all data must be correct, recorded in its entirety, and collected and recorded in the same manner throughout. Data must also be held in a manner that is lasting and accessible (Samson, 2021; U.S. Department of Health and Human Services et al., 2016).

Biopharmaceutical classification system (BCS). A system classifying a drug substance based on its minimum aqueous solubility in the pH range of 1–7.5, dose, and human fraction absorbed or intestinal membrane permeability. This system categorises drugs into four classes according to their permeability and solubility (Amidon et al., 1995). It has been suggested that the regulatory criterion for a highly soluble drug, whose highest dose (approved) strength is soluble in 250 mL aqueous media over the pH range of 1.0–6.8, is conservative for BCS Class I drugs and that further biowaivers for acidic drugs, BCS Class IIa, should be considered (Amidon et al., 1995).

Biorelevant performance classification system (BPCS). A system developed to (i) identify effective range of release achievable in population subsets and (ii) develop new release systems using models that self-learn from clinical outcomes and/or endpoints.

Chemistry, manufacture and control (CMC). Crucial activities when developing new pharmaceutical products. CMC involves defining manufacturing practices and product specifications that must be followed and met to ensure product safety and consistency between batches. CMC begins after a lead compound is identified through drug discovery and

continues through all remaining stages of the drug development life cycle. In addition to the pharmaceutical product, CMC also applies to the facility where manufacturing occurs.

Critical material attribute a.k.a. critical quality attributes of materials (CMA). A measurable material property whose variability has an impact on a critical quality attribute and therefore it should be monitored and controlled to ensure desired drug product quality.

Critical process parameter (CPP). A term used in pharmaceutical production for process variables which have an impact on a critical quality attribute (CQA) and, therefore, should be controlled to ensure the drug product obtains the desired quality (International Council for Harmonisation, 2009).

Critical quality attribute (CQA). A measurable physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality. It is primarily based upon the severity of harm and does not change as a result of risk management (International Council for Harmonisation, 2009).

Crystallisation classification system (CCS). A predictive classification system spanning the production of primary particles to formulated product and addressing manufacturability, stability, and performance parameters, which is being used to develop integrated platforms to support efficient and science-driven development from molecule to particle.

Cyber-physical research infrastructure (CPRI). UK Research and Innovation (UKRI) defines this as “the integration of digital and physical systems to create new capabilities and opportunities for research and innovation” (Simon Hart, 2023).

Cyber-physical system (CPS). The interlink of computational technologies with physical processes to analyse, monitor and/or control their functionality in a consistent, robust, safe, efficient and concurrent manner (Alguliyev et al., 2018; Baheti and Gill, 2011; Lee, 2006; Marwedel, 2021; Sanislav and Miclea, 2012).

Data fabric. A management system that enables full integration of a number of different data sources, pipelines and storage (IBM, 2024). This allows active metadata to be generated and enables collection of FAIR data. Data fabrics enable these data to also be accessible to humans, applications and other systems further down the supply chain (Blohm et al., 2024; Hechler et al., 2023).

Data lakes (DLs). Larger, less organised archives than data warehouses, which do not have a fixed structure (Dibley, 2022; Hlupić et al., 2022) and are used for storage and governance of a large range of data including structured and unstructured data (Hlupić et al., 2022; Nambiar and Mundra, 2022b). DLs are designed for decision-making and analysis, as they enable storage and processing in near real-time (Hlupić et al., 2022)

Data mesh (DM). A management system where data is collected by domain owners who generate data products. Multiple systems can then be combined and utilised by a range of users as required in a "shopping for data" and "self-service" manner, making reusable data more accessible (Blohm et al., 2024; Hechler et al., 2023).

Data warehouse (DW). A highly ordered archive that houses, organises and structures historical data. It enables data from a range of sources and geographical locations to be easily accessed either on premises or as a Cloud-based platform (Garani et al., 2019; Garcia et al., 2008; Nambiar and Mundra, 2022; Thantilage et al., 2023).

DataFactories. An autonomous experimental platform capable of collecting targeted experimental data for APIs, excipients, and products under a wide range of conditions exploiting automated dosing or sample handling, mobile robotics, small-scale experiments with integrated sensing/analytics/imaging for information extraction and global optimisation for self-learning experimental planning to meet objectives.

Design space. The combination of materials and process conditions which provide assurance of quality for a pharmaceutical product. This can be defined by determining the bounds of the critical process parameters and critical material attributes that guarantee the attainment of the targeted critical quality attributes (International Council for Harmonisation, 2009).

Developability classification system (DCS). A methodology of categorising a drug substance, building on the BCS, to account for the effects of an approximation of human fasted state intestinal solubility, a given solubility limited absorbable dose and a given dissolution rate in relation to particle size. This allows identification of development risks and enables CQAs to be identified for APIs exhibiting dissolution rate limited absorption (Butler and Dressman, 2010).

Digital-first. During development, *in silico* modelling is used to inform and guide process design before any experimental work is undertaken. The sole purpose of any initial experimental work is to achieve model parameterisation.

Digital twin (DT). Integrated digital framework to collate, analyse, visualise, and apply data, models, and knowledge of the rapid design, control, operation, and testing of continuous and modular processes for active pharmaceutical ingredient (API) crystallisation and drug product

(DP) production. The DT will combine the overarching digital definition of the materials, products, equipment, and processes. The FDA defines digital twins as a group of informational structures that simulate the configuration, framework and performance of a physical instrument or experiment and have a synergistic relationship with the physical twin by utilising live data from the physical twin, whilst informing physical next steps (U.S. Food & Drugs Administration, 2024).

Digital twin supply chain (DTSC). The above applied to the field of supply chain to allow simulation of the broader context of pharmaceutical processing.

Digitalisation. A process that can include both the increased use of robotics, automated solutions, and computerisation, thereby allowing reduced costs, improved efficiency and productivity, and increased flexibility.

Drug product (DP). A finished dosage form, e.g., tablet, capsule, or solution, which contains a drug substance, generally, but not necessarily, in association with one or more other ingredients.

Drug product performance. *In vivo* may be defined as the release of the drug substance from the drug product leading to bioavailability of the drug substance. The assessment of drug product performance is important since bioavailability is related both to the pharmacodynamic response and to adverse events.

Drug substance (DS). An active ingredient that is intended to furnish pharmacological activity or other direct effect in the diagnosis, cure, mitigation, treatment, or prevention of disease or to affect the structure or any function of the human body but does not include intermediates used in the synthesis of such ingredient.

Efficiency (of API production). Optimal operating parameters to meet sustainability and volumetric throughput per unit time targets.

EPSRC Future Manufacturing Research Hub in Continuous Manufacturing and Advanced Crystallisation (CMAC). A national centre for medicines manufacturing research, skills, technology and translation (CMAC, 2022).

Excipient. A constituent of a medicine other than the active substance, added in the formulation for a specific purpose (such as binding, disintegration or lubrication). While most excipients are considered inactive, some can have a known action or effect in certain

circumstances, which may enhance or control API performance (Kar et al., 2018; The International Pharmaceutical Excipients Council (IPEC Federation), 2023).

Extract-transform-load (ETL) or Extract-load-transform (ELT). ETL is the processing of data on a distinct server before sending to a data warehouse, whereas ELT involves processing of data within the data warehouse and thus raw, unstructured data can be sent directly to a data warehouse in this way, removing the need for a step-wise approach (Bartley, 2024).

Findable, accessible, interoperable, reusable (FAIR). Principles for good data practice, first established in 2016 (Wilkinson et al., 2016).

Generative pre-trained transformer (GPT) model. A large language model trained on unlabelled text which generates original human-like responses using an artificial neural network (Eloundou et al., 2023).

Installation qualification (IQ). In the context of QbDD, this can be considered as documentation of the model having met the requirements defined by the model designer for configuration and initial implementation, possibly defined by an installation checklist, system specifications and/or datasheets (Egnyte, 2024; The FDA Group LLC, 2024). This can include suitable use, associated software requirements, environmental requirements, and calibration and verification requirements (Precision Solutions Inc., 2024).

Machine learning (ML). Computer-based development of algorithms for problem solving where the computer can learn and adapt without human interaction that may be used to train AI (U.S. Food & Drugs Administration, 2024).

Manufacturability (of drug product). The properties of a drug substance to be manufactured by an intended route for a desired formulation.

Manufacturing classification system (MCS). A means of categorising drug products based on processing route. It summarises conclusions from a dedicated Academy of Pharmaceutical Sciences (APS) conference and subsequent discussion within APS focus groups and the MCS working party (Leane et al., 2018, 2015). The MCS is intended as a tool for pharmaceutical scientists to rank the feasibility of different processing routes for the manufacture of oral solid dosage forms, based on selected properties of the API and the needs of the formulation (Leane et al., 2024).

Manufacturing classification system+ (MCS+). A system developed by the International Society for Pharmaceutical Engineering (ISPE) (Potter, 2022), that provides a framework for classifying manufacturing processes based on their complexity and potential impact on product quality. MCS+ builds upon the original Manufacturing Classification System (MCS) developed by ISPE, but includes additional factors such as process variability, criticality of process steps, and complexity of equipment and automation (Potter, 2022).

Model Validation. In modelling (and more specifically within this publication), this is evaluation of the model outputs against an independent data set that has known outputs and has not been used in the training of the model.

Objective. A quantitative or qualitative value or goal for which the achievement thereof defines the success of an optimisation, be the optimisation machine learning or otherwise.

Operation qualification (OQ). In the context of QbDD, this consists of establishing and assessing the various aspects of the model which may affect the overall quality of the process controlled (The FDA Group LLC, 2024). It ensures reproducibility and reliability within appropriate operating conditions and that strategies for maintenance, deviation checks, performance checks and calibrations are put in place (Powder Systems, 2024; Precision Solutions Inc., 2024).

Performance qualification (PQ). A verification step for equipment or model use in which a qualification and verification group monitors, checks and reports if the quality requirements are achieved, ensuring reliability over time (Powder Systems, 2024; The FDA Group LLC, 2024). Methodologies and validations to this end could include the following: data summaries, suitable calibrations and validations, variability limits and experimental verification strategies (Precision Solutions Inc., 2024).

Process analytical technology (PAT). Mechanism to design, analyse, and control pharmaceutical manufacturing processes through measurement of material and quality attributes (U.S. Department of Health and Human Services et al., 2004).

Process Validation. The FDA defines this as “the collection and evaluation of data which establishes scientific evidence that a process is capable of consistently delivering quality product throughout the product lifecycle” (Bizjak and U.S. Food & Drug Administration, 2020; Tartal and U.S. Food & Drug Administration, 2015; U.S. Department of Health and Human Services et al., 2011). It generally consist of 3 stages:

1. Initial R&D experimentation and risk assessment to give an indication of the nature of the manufacturing process, allowing development of a control strategy.

2. IQ, OQ and PQ to ascertain the suitability of the technology utilised for its proposed purpose. Planning and carrying out of experimentation to prove this.

3. Continued observation of the process and technology utilised to allow optimisation based on data collected and human experience (Bizjak and U.S. Food & Drug Administration, 2020; U.S. Department of Health and Human Services et al., 2011).

Process Verification. The FDA defines this as “confirmation by examination and provision of objective evidence that specified requirements have been fulfilled” (Tartal and U.S. Food & Drug Administration, 2015). This is vital to ensure a controlled state is maintained throughout a process (U.S. Department of Health and Human Services et al., 2011).

Quality. The suitability of either a drug substance or a drug product for its intended use. Includes attributes such as identity, strength, and purity.

Quality by control (QbC). Also referred to as quality control (QC). Control strategy where active process control ensures product quality. This builds on QbT by implementation of PAT and seeks to rectify issues with a lack of integration between unit operations. It can be considered as the proposal and implementation of a manufacturing system using an active process control system developed in agreement with process automation principles, dictated by a strong degree of quantitative and predictive product and process understanding (Su et al., 2019).

Quality by design (QbD). An efficient development procedure that commences with predetermined objectives and focuses on product and process understanding and control with a strong basis in rigorous science and quality risk management (International Council for Harmonisation, 2009; Yu et al., 2014)

Quality by digital design (QbDD). Application of extensive modelling and data driven decision support tools to quickly, robustly and sustainably drive an efficient development procedure that commences with predetermined objectives and focuses on product and process understanding and control with a strong basis in rigorous science and quality risk management (International Council for Harmonisation, 2009; Yu et al., 2014)

Quality by testing (QbT). Also referred to as quality assurance (QA). A traditional control strategy based on batch processing, which assesses the quality of a manufactured medicine by testing the final product to determine if targets have been met (Rege et al., 2024; Yu et al., 2014).

Quality management system. A structured framework that ensures that a medicine manufacturer consistently meets patient requirements and regulatory standards. It includes policies, procedures, processes and resources that guide quality-related activities.

Quality target product profile (QTPP). A prospective summary of the quality characteristics of a drug product that must be achieved to ensure the desired quality, taking into account safety and efficacy of the drug product.

Sensitivity analysis. A technique to establish how a range of values of an independent variable influence a dependent variable within a particular hypothesis, or effectively how causes of uncertainty within a model impact its overarching uncertainty.

Supply chain resilience. The ability to anticipate and/or act on disruptions, to achieve a rapid and economical recovery, and thus regain the normal running of operations (Tukamuhabwa et al., 2015).

Sustainability. Process targets to drive reduction in materials, energy, resources, carbon footprint and environmental impact etc. Often set at organisational level.

Uncertainty. In modelling, this is considered a quantity that enables modellers to assess the accuracy and reliability of models and to make informed decisions based on the results.

Workflow. Systematic, science-based process design sequence of tasks. Experimental, computational, and analytical tasks should be clearly defined.

5. Acknowledgments

This work was supported by the EPSRC Future Continuous Manufacturing and Advanced Crystallisation Research Hub (grant ref: EP/P006965/1) with contributions from Digital Medicines Manufacturing (DM2) Research Centre co-funded by the Made Smarter Innovation challenge at UK Research and Innovation (grant ref: EP/V062077/1), GIBio (Gastrointestinal bioreactor to evaluate ingestible medicines and inform formulation and manufacture) (grant ref: EP/W036452/1), a Knowledge Transfer Partnership with Siemens Industry Software Limited (partnership number: 11937) and the CMAC National Facility. The authors would also like to thank Steven Scott and his colleagues at TwoFifths for their support in producing the artwork associated with this publication.

6. Electronic Supplementary Information

ESI 1: Evidence of digital adoption by regulatory bodies.

Source	Initiative	Purpose	Reference(s)

U.S. Food & Drug Administration (FDA)	Technology Modernisation Action Plan	Set out FDA's plans for future technical infrastructure and their continuing data management procedures, with particular regards to collaboration between stakeholders	(U.S. Food & Drug Administration, 2019)
	Data Modernisation Action Plan	Identified the need for high-value driver projects with robust and harmonious data collection and management throughout the regulatory body, while upskilling and developing a talented workforce to this end	(U.S. Food & Drug Administration, 2021)
	Knowledge-aided Assessment & Structured Application (KASA)	• Enables collection and management of knowledge throughout product development. • Generates rules and algorithms for risk assessment, control and communication. • Enables IT assisted analysis of applications to assess regulatory standards and quality risks. • Generates a strategic assessment	(Yu et al., 2019)
	Report on the use of modelling and simulation, 2022	• The use of modelling and simulation a vital regulatory science resource • The formation of a Modeling and Simulation Working Group	(Riaz Ahmed et al., 2022)
	Center for Drug Evaluation and Research: Artificial Intelligence in Drug Manufacturing	Sets out: • The use of AI in pharmaceutical manufacturing • Gathering information on public opinion on use of AI	(U.S. Food & Drug Administration, 2023a)
	Using Artificial Intelligence & Machine Learning in the Development of Drug and Biological Products	Sets out: • Present and possible future uses of AI and ML • Reflections on AI and ML with regards to thoughts on risks and concerns about implementation • Stakeholder engagement	(U.S. Food & Drug Administration, 2024a)
	Digital Health Center of Excellence	Centre aiming to: • Encourage collaboration • Encourage knowledge exchange	(U.S. Food & Drug Administration, 2024b)

		• Transform regulatory procedures	
	Emerging Technology Program (ETP)	Set out Emerging Technology Team (ETT) aims to work with industry to help guide partners in their new technologically advanced regulatory submissions.	(U.S. Food & Drug Administration, 2023b)
	FDA/PQRI Workshop on the Regulatory Framework for the Utilization of Artificial Intelligence in Pharmaceutical Manufacturing	Workshop aiming to encourage collaboration between AI stakeholders on R&D, introduction and regulatory submission in various areas such as process development and control, quality management, lifecycle management, and GMP.	(U.S. Food & Drug Administration, 2023c)
	Real-World Evidence	Set out the use of live data to provide information on lifecycle DP safety	(U.S. Food & Drug Administration, 2024c)
	Digital Health Technologies (DHTs) for Drug Development	Sets out DHT advantages including: • Real time monitoring and identification of deviations • Collect and analyse deviations that would not be picked up during routine visits • Promotion of remote data collection	(U.S. Food & Drug Administration, 2024d)
	Artificial Intelligence and Machine Learning (AI/ML) for Drug Development	Sets out FDA stance on AI and ML use	(U.S. Food & Drug Administration, 2024e)
European Medicines Agency (EMA)	Information Management Strategy, 2015	To standardise processes and inform regulation in the rest of the world	(Pauli and Williams, 2018)
	EMA Network Strategy to 2025	• Identification of six core sections including data analytics, digital tools and digital transformation • Assessment and analysis of healthcare and clinical data, generating EU network potential to analyse large data sets, encouraging development of novel regulatory practices, promoting cyber physical technologies and encouraging	(European Medicines Agency and Heads of Medicines Agencies, 2020)

		secure data management practices	
	Electronic format for product information by adoption of the International Standardisation Organisation Identification and Description of Medicinal and Pharmaceutical Products (ISO IDMP) standards	Consistent exchange of medicinal product information in a reliable and consistent fashion, by generating a common product 'language' for stakeholders	(European Medicines Agency, 2016)
	Joint Biologics Working Party / Quality Working Party workshop with stakeholders in relation to prior knowledge and its use in regulatory applications.	The definition of the use of prior knowledge within innovative methods	(European Medicines Agency, 2017)
	New online platform for scientific advice	New regulatory submission portal (IRIS Platform)	(European Medical Agency, 2020)
Insilico UK, Medicines & Healthcare Products Regulatory Agency, Royal Academy of Engineering and PHG Foundation	Cross-Regulator Workshop: Journeys, experiences and best practices on computer modelled and simulated regulatory evidence—Workshop Report.	Establishes potential for modelling and simulation in improving patient safety and cost-effectiveness within the life sciences sector; Discussion of issues limiting adoption and, need for standardisation and quality control	(Redrup Hill et al., 2023)

7. References

Aamir, E., 2010. Population Balance Model-Based Optimal Control of Batch Crystallisation Processes for Systematic Crystal Size Distribution Design (Doctor of Philosophy). Loughborough University, Loughborough.

Abasi, N., Sohrabi, M.R., Motiee, F., Davallo, M., 2021. The application of artificial neural network and least square support vector machine methods based on spectrophotometry method for the rapid simultaneous estimation of triamcinolone, neomycin, and nystatin in skin ointment formulation. Optik (Stuttg) 241. https://doi.org/10.1016/j.ijleo.2021.167210

Abbas, F., Salehian, M., Hou, P., Moores, J., Goldie, J., Tsioutsios, A., Portela, V., Boulay, Q., Thiolliere, R., Stark, A., Schwartz, J.-J., Guerin, J., Maloney, A.G.P., Moldovan, A.A., Reynolds, G.K., Mantanus, J., Clark, C., Chapman, P., Florence, A., Markl, D., 2025. Accelerated Medicines Development using a Digital Formulator and a Self-Driving Tableting DataFactory. https://doi.org/10.48550/arXiv.2503.17411

Abouzied, A.S., Alshahrani, S.M., Hani, U., Obaidullah, A.J., Al Awadh, A.A., Lahiq, A.A., Al-fanhrawi, H.J., 2023. Assessment of solid-dosage drug nanonization by theoretical advanced models: Modeling of solubility variations using hybrid machine learning models. Case Studies in Thermal Engineering 47. https://doi.org/10.1016/j.csite.2023.103101

Abuhassan, Q., Khadra, I., Pyper, K., Augustijns, P., Brouwers, J., Halbert, G.W., 2022a. Structured solubility behaviour in bioequivalent fasted simulated intestinal fluids. European Journal of Pharmaceutics and Biopharmaceutics 176, 108–121. https://doi.org/10.1016/j.ejpb.2022.05.010

Abuhassan, Q., Khadra, I., Pyper, K., Augustijns, P., Brouwers, J., Halbert, G.W., 2022b. Fasted intestinal solubility limits and distributions applied to the biopharmaceutics and developability classification systems. European Journal of Pharmaceutics and Biopharmaceutics 170, 160–169. https://doi.org/10.1016/j.ejpb.2021.12.006

Abuhassan, Q., Silva, M.I., Tamimi, R.A.R., Khadra, I., Batchelor, H.K., Pyper, K., Halbert, G.W., 2024. A novel simulated media system for in vitro evaluation of bioequivalent intestinal drug solubility. European Journal of Pharmaceutics and Biopharmaceutics 199. https://doi.org/10.1016/j.ejpb.2024.114302

Acceleration Consortium, University of Toronto, 2024. Intrepid: build better medicines, fast [WWW Document]. Acceleration Consortium. URL https://acceleration.utoronto.ca/news/intrepid-build-better-medicines-fast (accessed 7.29.24).

Adabala Viswa, C., Zurkiya, D., Bleys, J., Shah, B., Leydon, E., 2024. Generative AI in the pharmaceutical industry: Moving from hype to reality [WWW Document]. McKinsey & Company. URL https://www.mckinsey.com/industries/life-sciences/our-insights/generative-ai-in-the-pharmaceutical-industry-moving-from-hype-to-reality (accessed 4.2.24).

Agrawal, P., Rawal, S.H., Reddy, V.R., Viswanath, S.K., Merritt, J.M., 2023. Case Studies in the Application of a Workflow-Based Crystallization Design for Optimized Impurity Rejection in Pharmaceutical Development. Org Process Res Dev 27, 610–626. https://doi.org/10.1021/acs.oprd.2c00346Ahluwalia, K., Abernathy, M.J., Beierle, J., Cauchon, N.S., Cronin, D., Gaiki, S., Lennard, A., Mady, P., McGorry, M., Sugrue-Richards, K., Xue, G., 2022. The Future of CMC Regulatory Submissions: Streamlining Activities Using Structured Content and Data Management. J Pharm Sci. https://doi.org/10.1016/j.xphs.2021.09.046

Ahmed, S., Burgos, R., Roy, C., Boroyevich, D., Mattavelli, P., Wang, F., 2012. Modeling Verification, Validation, and Uncertainty Quantification (VV&UQ) Procedure for a Two-Level Three-Phase Boost Rectifier, in: Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC). Orlando, FL, USA, pp. 1894–1901. https://doi.org/10.1109/APEC.2012.6166081

Akundi, A., Euresti, D., Luna, S., Ankobiah, W., Lopes, A., Edinbarough, I., 2022. State of Industry 5.0—Analysis and Identification of Current Research Trends. Applied System Innovation 5. https://doi.org/10.3390/asi5010027

AlAlaween, W.H., Mahfouf, M., Salman, A.D., 2016. Predictive modelling of the granulation process using a systems-engineering approach. Powder Technol 302, 265–274. https://doi.org/10.1016/j.powtec.2016.08.049Alguliyev, R., Imamverdiyev, Y., Sukhostat, L., 2018. Cyber-physical systems and their security issues. Comput Ind. https://doi.org/10.1016/j.compind.2018.04.017

Allain, L., Baertschi, S.W., Clapham, D., Foti, C., Lantaff, W.M., Reed, R.A., Templeton, A.C., Tønnesen, H.H., 2016. Implications of In-Use Photostability: Proposed Guidance for Photostability Testing and Labeling to Support the Administration of Photosensitive Pharmaceutical Products, Part 3. Oral Drug Products. J Pharm Sci 105, 1586–1594. https://doi.org/10.1016/j.xphs.2016.02.035

Amidon, G.L., Lennernäs, H., Shah, V.P., Crison, J.R., 1995. A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists. https://doi.org/10.1023/A:1016212804288

An, G., Cockrell, C., 2022. Drug Development Digital Twins for Drug Discovery, Testing and Repurposing: A Schema for Requirements and Development. Frontiers in Systems Biology 2. https://doi.org/10.3389/fsysb.2022.928387

Arden, N.S., Fisher, A.C., Tyner, K., Yu, L.X., Lee, S.L., Kopcha, M., 2021. Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. Int J Pharm. https://doi.org/10.1016/j.ijpharm.2021.120554

Aspuru-Guzik Group, 2024. AI for Discovery and Self-Driving Labs [WWW Document]. The Matter Lab. URL https://www.matter.toronto.edu/basic-content-page/ai-for-discovery-and-self-driving-labs (accessed 7.29.24).

Axelrod, C.W., 2013. Managing the risks of cyber-physical systems, in: 2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT). Presented at the 2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT), IEEE, Farmingdale, NY, USA, pp. 1–6. https://doi.org/10.1109/LISAT.2013.6578215

Azad, M.A., Capellades, G., Wang, A.B., Klee, D.M., Hammersmith, G., Rapp, K., Brancazio, D., Myerson, A.S., 2021. Impact of Critical Material Attributes (CMAs)-Particle Shape on Miniature Pharmaceutical Unit Operations. AAPS PharmSciTech 22. https://doi.org/10.1208/s12249-020-01915-6

Azizi, A., Vatankhah, R., Editors, B., 2023. Industry 4.0: Technologies, Applications, and Challenges. Emerging Trends in Mechatronics.

Bahaghighat, M., Akbari, L., Xin, Q., 2019. A machine learning-based approach for counting blister cards within drug packages. IEEE Access 7, 83785–83796. https://doi.org/10.1109/ACCESS.2019.2924445

Baheti, R., Gill, H., 2011. Cyber-physical Systems, in: T, S., A.M, A. (Eds.), The Impact of Control Technology.

Bai, J., Mosbach, S., Taylor, C.J., Karan, D., Lee, K.F., Rihm, S.D., Akroyd, J., Lapkin, A.A., Kraft, M., 2024. A dynamic knowledge graph approach to distributed self-driving laboratories. Nat Commun 15. https://doi.org/10.1038/s41467-023-44599-9

Bala, I., Verma, S., Namdev, K.K., 2014. A Review on Process Variables Effects of an Oral Dosage Form by Using “Quality by Design,” Available online on www.ijddt.com International Journal of Drug Delivery Technology.

Barrasso, D., El Hagrasy, A., Litster, J.D., Ramachandran, R., 2015. Multi-dimensional population balance model development and validation for a twin screw granulation process. Powder Technol 270, 612–621. https://doi.org/10.1016/j.powtec.2014.06.035

Barshikar, R., 2019. Quality by Design (QbD) and its implementation in Pharma Industry [WWW Document]. Express Pharma. URL https://www.expresspharma.in/quality-by-design-qbd-and-its-implementation-in-pharma-industry/ (accessed 4.9.24).

Bartley, K., 2024. ETL vs ELT: What's The Difference? [WWW Document]. Rivery. URL https://rivery.io/blog/etl-vs-elt/ (accessed 7.27.24).Beckmann, K., 2023. How ChatGPT & Co. Will Change Our Work: Four Hypotheses About Working with (Generative) Artificial Intelligence [WWW Document]. Merck Group. URL https://www.merckgroup.com/en/the-future-transformation/how-chatgpt-co-change-our-work-four-hypotheses-working.html (accessed 9.4.23).

Bekaert, B., Van Snick, B., Pandelaere, K., Dhondt, J., Di Pretoro, G., De Beer, T., Vervaet, C., Vanhoorne, V., 2022. Continuous direct compression: Development of an empirical predictive model and challenges regarding PAT implementation. Int J Pharm X 4. https://doi.org/10.1016/j.ijpx.2021.100110

Bellinghausen, S., 2020. Modelling and scaling rules for high-shear wet granulation of pharmaceuticals. The University of Sheffield, Sheffield.Berkeley Lab, 2024. The Materials Project [WWW Document]. Lawrence Berkeley National Laboratory, The University of Berkeley. URL https://next-gen.materialsproject.org/ (accessed 7.29.24).

Bermingham, B., 2023. Digital Transformation: The best foot forward for pharma [WWW Document]. Intuition. URL https://www.intuition.com/digital-transformation-the-best-foot-forward-for-pharma/#:~:text=Digital%20transformation%20offers%20pharmaceutical%20organizations,facing%20services%20to%20run%20smoother. (accessed 11.4.24).

Bhalode, P., Chen, Y., Ierapetritou, M., 2022. Hybrid Modelling Strategies for Continuous Pharmaceutical Manufacturing within Digital Twin Framework, in: Computer Aided Chemical Engineering. Elsevier B.V., pp. 2125–2130. https://doi.org/10.1016/B978-0-323-85159-6.50354-7

Bhardwaj, R.M., Johnston, A., Johnston, B.F., Florence, A.J., 2015. A random forest model for predicting the crystallisability of organic molecules. CrystEngComm 17, 4272–4275. https://doi.org/10.1039/c4ce02403f

Biron, L., 2023. Meet the Autonomous Lab of the Future [WWW Document]. Lawrence Berkeley National Laboratory, The University of Berkeley. URL https://newscenter.lbl.gov/2023/04/17/meet-the-autonomous-lab-of-the-future/ (accessed 9.10.24).

Bizjak, T.G., U.S. Food & Drug Administration, 2020. CGMP and Process Validation, in: Compounding Quality Center of Excellence Virtual Conference: Working Together for Patient Safety. Virtual Conference.

Blanco-González, A., Cabezón, A., Seco-González, A., Conde-Torres, D., Antelo-Riveiro, P., Piñeiro, Á., Garcia-Fandino, R., 2023. The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies. Pharmaceuticals. https://doi.org/10.3390/ph16060891Blohm, I., Wortmann, F., Legner, C., Köbler, F., 2024. Data products, data mesh, and data fabric: New paradigm(s) for data and analytics? Business and Information Systems Engineering. https://doi.org/10.1007/s12599-024-00876-5

Boiko, D.A., MacKnight, R., Kline, B., Gomes, G., 2023. Autonomous chemical research with large language models. Nature 624, 570–578. https://doi.org/10.1038/s41586-023-06792-0Bolbot, V., Theotokatos, G., Bujorianu, L.M., Boulougouris, E., Vassalos, D., 2019.

Vulnerabilities and safety assurance methods in Cyber-Physical Systems: A comprehensive review. Reliab Eng Syst Saf. https://doi.org/10.1016/j.ress.2018.09.004

Boobier, S., Hose, D.R.J., Blacker, A.J., Nguyen, B.N., 2020. Machine learning with physicochemical relationships: solubility prediction in organic solvents and water. Nat Commun 11. https://doi.org/10.1038/s41467-020-19594-zBordukova, M., Makarov, N., Rodriguez-Esteban, R., Schmich, F., Menden, M.P., 2024. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials. Expert Opin Drug Discov. https://doi.org/10.1080/17460441.2023.2273839

Bran, A.M., Cox, S., Schilter, O., Baldassari, C., White, A.D., Schwaller, P., 2024. Augmenting large language models with chemistry tools. Nature Machine Intelligence 6, 525–535. https://doi.org/10.1038/s42256-024-00832-8

Brown, C.J., Mcglone, T., Yerdelen, S., Srirambhatla, V., Mabbott, F., Gurung, R., Briuglia, M.L., Ahmed, B., Polyzois, H., Mcginty, J., Perciballi, F., Fysikopoulos, D., Macfhionnghaile, P., Siddique, H., Raval, V., Harrington, T.S., Vassileiou, A.D., Robertson, M., Prasad, E., Johnston, A., Johnston, B., Nordon, A., Srai, J.S., Halbert, G., Ter Horst, J.H., Price, C.J., Rielly, C.D., Sefcik, J., Florence, A.J., 2018. Enabling precision manufacturing of active pharmaceutical ingredients: workflow for seeded cooling continuous crystallisations. Mol Syst Des Eng 3, 518–549. https://doi.org/10.1039/c7me00096k

Brownett-Gale, S., 2024. Has AI Revolutionised Pharmaceutical Supply Chain Management? Pharmaceutical Packaging Expert Weighs In [WWW Document]. GHP. URL https://ghpnews.digital/has-ai-revolutionised-pharmaceutical-supply-chain-management-pharmaceutical-packaging-expert-weighs-in/ (accessed 9.23.24).

Burger, B., Maffettone, P.M., Gusev, V. V., Aitchison, C.M., Bai, Y., Wang, X., Li, X., Alston, B.M., Li, B., Clowes, R., Rankin, N., Harris, B., Sprick, R.S., Cooper, A.I., 2020. A mobile robotic chemist. Nature 583, 237–241. https://doi.org/10.1038/s41586-020-2442-2

Butler, J.M., Dressman, J.B., 2010. The developability classification system: Application of biopharmaceutics concepts to formulation development. J Pharm Sci 99, 4940–4954. https://doi.org/10.1002/jps.22217

Butters, M., Catterick, D., Craig, A., Curzons, A., Dale, D., Gillmore, A., Green, S.P., Marziano, I., Sherlock, J.P., White, W., 2006. Critical assessment of pharmaceutical processes - A rationale for changing the synthetic route. Chem Rev. https://doi.org/10.1021/cr050982w

Candelon, F., Krayer, L., Rajendran, S., Zhukov, L., 2023. Generative A.I.: How to future-proof your company's strategy [WWW Document]. Fortune. URL https://fortune.com/2023/07/07/generative-ai-strategy-4-things-executives-should-do-bcg/ (accessed 9.4.23).

Caphart, M., Sausville, R., Liang, J., Maroney Benassi, P., 2006. Guidance for Industry on Quality Systems Approach to Pharmaceutical Current Good Manufacturing Practice Regulations [WWW Document]. National Archives: Federal Register . URL https://www.federalregister.gov/documents/2006/10/02/E6-16215/guidance-for-industry-on-quality-systems-approach-to-pharmaceutical-current-good-manufacturing#:~:text=The%20Food%20and%20Drug%20Administration%20(FDA)%20is%20announcing%20the%20availability,in%20the%20quality%20and%20manufacturing (accessed 4.5.24).Carou-Senra, P., Ong, J.J., Castro, B.M., Seoane-Viaño, I., Rodríguez-Pombo, L., Cabalar, P., Alvarez-Lorenzo, C., Basit, A.W., Pérez, G., Goyanes, A., 2023. Predicting pharmaceutical inkjet printing outcomes using machine learning. Int J Pharm X 5. https://doi.org/10.1016/j.ijpx.2023.100181

Certara, 2023. Simcyp PBPK [WWW Document]. Certara. URL https://www.certara.com/software/simcyp-pbpk/ (accessed 9.19.23).

Chakravarty, K., Antontsev, V., Bundey, Y., Varshney, J., 2021. Driving success in personalized medicine through AI-enabled computational modeling. Drug Discov Today. https://doi.org/10.1016/j.drudis.2021.02.007

Chalortham, N., Ruangrajitpakorn, T., Supnithi, T., Leesawat, P., 2013. OXPIRT: Ontology-based eXpert system for Production of a generic Immediate Release Tablet, in: Formulation Tools for Pharmaceutical Development. Elsevier Ltd, pp. 203–228. https://doi.org/10.1533/9781908818508.203

Chaudhury, A., Wu, H., Khan, M., Ramachandran, R., 2014. A mechanistic population balance model for granulation processes : Effect of process and formulation parameters. Chem Eng Sci 107, 76–92. https://doi.org/10.1016/j.ces.2013.11.031

Chemical Entities of Biological Interest (ChEBI), 2024. ChEBI. Chemical Entities of Biological Interest (ChEBI).Chen, C., 2024. Artificial Intelligence aided pharmaceutical engineering: Development of hybrid machine learning models for prediction of nanomedicine solubility in supercritical solvent. J Mol Liq 397. https://doi.org/10.1016/j.molliq.2024.124127

Chen, K., Li, J., Wang, K., Du, Y., Yu, J., Lu, J., Li, L., Qiu, J., Pan, J., Huang, Y., Fang, Q., Heng, P.A., Chen, G., 2023. Chemist-X: Large Language Model-empowered Agent for Reaction Condition Recommendation in Chemical Synthesis.

Chen, Y., Li, Y., 2003. A new model for predicting moisture uptake by packaged solid pharmaceuticals. Int J Pharm 255, 217–225. https://doi.org/10.1016/S0378-5173(03)00089-9

Chen, Y., Sampat, C., Huang, Y.S., Ganesh, S., Singh, R., Ramachandran, R., Reklaitis, G. V., Ierapetritou, M., 2023. An integrated data management and informatics framework for continuous drug product manufacturing processes: A case study on two pilot plants. Int J Pharm 642. https://doi.org/10.1016/j.ijpharm.2023.123086

Christensen, M., Yunker, L.P.E., Adedeji, F., Häse, F., Roch, L.M., Gensch, T., dos Passos Gomes, G., Zepel, T., Sigman, M.S., Aspuru-Guzik, A., Hein, J.E., 2021. Data-science driven autonomous process optimization. Commun Chem 4. https://doi.org/10.1038/s42004-021-00550-xCMAC, 2022. EPSRC Future Manufacturing Hub in Continuous Manufacturing and Advanced Crystallisation [WWW Document]. URL https://cmac.ac.uk/ (accessed 9.6.23).

CMAC, 2021. CMAC Strategy 2021-2021. EPSRC Future Manufacturing Research Hub.

Coley, C.W., Green, W.H., Jensen, K.F., 2018. Machine Learning in Computer-Aided Synthesis Planning. Acc Chem Res 51, 1281–1289. https://doi.org/10.1021/acs.accounts.8b00087

Coley, C.W., Thomas, D.A., Lummiss, J.A.M., Jaworski, J.N., Breen, C.P., Schultz, V., Hart, T., Fishman, J.S., Rogers, L., Gao, H., Hicklin, R.W., Plehiers, P.P., Byington, J., Piotti, J.S., Green, W.H., John Hart, A., Jamison, T.F., Jensen, K.F., 2019. A robotic platform for flow synthesis of organic compounds informed by AI planning. Science (1979) 365. https://doi.org/10.1126/science.aax1566Conroy, D., 2023. Accelerating Drug Development To Keep Pace With Drug Discovery [WWW Document]. Forbes. URL https://www.forbes.com/sites/forbestechcouncil/2023/03/10/accelerating-drug-development-to-keep-pace-with-drug-discovery/ (accessed 3.27.24).

Cook, J., Cruañes, M.T., Gupta, M., Riley, S., Crison, J., 2014. Quality-by-design: Are we there yet? AAPS PharmSciTech 15, 140–148. https://doi.org/10.1208/s12249-013-0043-1

Cramer, J.A., 1998. Enhancing Patient Compliance in the Elderly: Role of Packaging Aids and Monitoring. Drugs Aging 12, 7–15.

Creaner, G., Fitzgerald, D., 2024. What is IQ OQ PQ and why are they critical in the Pharmaceutical Industry? [WWW Document]. Get Reskilled. URL https://www.getreskilled.com/what-are-iq-oq-pq/ (accessed 5.30.24).

Crews, C., Kenny, P.S., O'Flynn, D., Speller, R.D., 2018. Multivariate calibration of energy-dispersive X-ray diffraction data for predicting the composition of pharmaceutical tablets in packaging. J Pharm Biomed Anal 151, 186–193. https://doi.org/10.1016/j.jpba.2017.12.036

Cumming, J.G., Davis, A.M., Muresan, S., Haeberlein, M., Chen, H., 2013. Chemical predictive modelling to improve compound quality. Nat Rev Drug Discov. https://doi.org/10.1038/nrd4128

Dai, S., Xu, B., Zhang, Z., Yu, J., Wang, F., Shi, X., Qiao, Y., 2019. A compression behavior classification system of pharmaceutical powders for accelerating direct compression tablet formulation design. Int J Pharm 572. https://doi.org/10.1016/j.ijpharm.2019.118742

Dara, S., Dhamercherla, S., Jadav, S.S., Babu, C.M., Ahsan, M.J., 2022. Machine Learning in Drug Discovery: A Review. Artif Intell Rev 55, 1947–1999. https://doi.org/10.1007/s10462-021-10058-4

David, L., Thakkar, A., Mercado, R., Engkvist, O., 2020. Molecular representations in AI-driven drug discovery: a review and practical guide. J Cheminform. https://doi.org/10.1186/s13321-020-00460-5

Davis, B., Schlindwein, W.S., 2018. Introduction to Quality by Design (QbD), in: Pharmaceutical Quality by Design. Wiley, pp. 1–9. https://doi.org/10.1002/9781118895238.ch1

Department of Health and Human Services U.S. Food and Drug Administration, 2007. Pharmaceutical Quality for the 21st Century A Risk-Based Approach Progress Report [WWW Document]. U.S. Food and Drug Administration. URL https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/pharmaceutical-quality-21st-century-risk-based-approach-progress-report (accessed 4.9.24).

Department of Health and Human Services U.S Food and Drug Administration, 2004. Pharmaceutical CGMPs for the 21st Century-A Risk-Based Approach: Final Report.

Deshpande, A., 2023. Machine Learning for Modeling and Control of a Packaging Manufacturing Process.Destro, F., Barolo, M., 2022. A review on the modernization of pharmaceutical development and manufacturing – Trends, perspectives, and the role of mathematical modeling. Int J Pharm. https://doi.org/10.1016/j.ijpharm.2022.121715Destro, F., Hur, I., Wang, V., Abdi, M., Feng, X., Wood, E., Coleman, S., Firth, P., Barton, A., Barolo, M., Nagy, Z.K., 2021. Mathematical modeling and digital design of an intensified filtration-washing-drying unit for pharmaceutical continuous manufacturing. Chem Eng Sci 244. https://doi.org/10.1016/j.ces.2021.116803Devogelaer, J.-J., Meekes, H., Tinnemans, P., Vlieg, E., De Gelder, R., 2020. Co-crystal Prediction by Artificial Neural Networks**. Angewandte Chemie 59, 21711–21718. https://doi.org/10.26434/chemrxiv.12452849.v2Dibley, J., 2022. The Modern Data Repository: Understanding Your Options [WWW Document]. Netwrix. URL

https://blog.netwrix.com/2022/12/09/the-modern-data-repository-understanding-your-options/ (accessed 7.30.24).

Dong, J., Wu, Z., Xu, H., Ouyang, D., 2024. FormulationAI: a novel web-based platform for drug formulation design driven by artificial intelligence. Brief Bioinform 25. https://doi.org/10.1093/bib/bbad419Dowden, H., Munro, J., 2019. Trends in clinical success rates and therapeutic focus. Nat Rev Drug Discov. https://doi.org/10.1038/d41573-019-00074-z

Durá, M., Sánchez-García, Á., Sáez, C., Leal, F., Chis, A.E., González-Vélez, H., García-Gómez, J.M., 2022. Towards a Computational Approach for the Assessment of Compliance of ALCOA+ Principles in Pharma Industry, in: Studies in Health Technology and Informatics. IOS Press BV, pp. 755–759. https://doi.org/10.3233/SHTI220578

Egnyte, 2024. A Guide to IQ, OQ, and PQ in FDA-Regulated Industries [WWW Document]. Egnyte. URL https://www.egnyte.com/guides/life-sciences/oq (accessed 5.30.24).Elbadawi, M., Gaisford, S., Basit, A.W., 2021. Advanced machine-learning techniques in drug discovery. Drug Discov Today. https://doi.org/10.1016/j.drudis.2020.12.003

Eloundou, T., Manning, S., Mishkin, P., Rock, D., 2023. GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models.

European Commission, 2021a. Industry 5.0 Towards a sustainable, human-centric and resilient European industry. Directorate-General for Research and Innovation. https://doi.org/10.2777/308407

European Commission, 2021b. Industry 5.0: A Transformative Vision for Europe ESIR Policy Brief No.3. Directorate-General for Research and Innovation. https://doi.org/10.2777/17322

European Commission, 2020. Enabling Technologies for Industry 5.0 Results of a workshop with Europe's technology leaders. https://doi.org/10.2777/082634

European Commission, Directorate-General for Research and Innovation, 2022. Industry 5.0 [WWW Document]. European Commission. URL https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/industry-50_en (accessed 9.9.24).

European Medical Agency, 2020. New online platform for scientific advice [WWW Document]. European Medical Agency. URL https://www.ema.europa.eu/en/news/new-online-platform-scientific-advice (accessed 4.26.24).

European Medicines Agency, 2024. Preliminary QIG Considerations regarding Pharmaceutical Process Models.

European Medicines Agency, 2017. Joint Biologics Working Party / Quality Working Party workshop with stakeholders in relation to prior knowledge and its use in regulatory applications [WWW Document]. European Medicines Agency. URL https://www.ema.europa.eu/en/events/joint-biologics-working-party-quality-working-party-workshop-stakeholders-relation-prior-knowledge-its-use-regulatory-applications (accessed 4.26.24).

European Medicines Agency, 2016. Introduction to ISO Identification of Medicinal Products, SPOR programme.European Medicines Agency, Heads of Medicines Agencies, 2020. European medicines agencies network strategy to 2025 - Protecting public health at a time of rapid change.Feenstra, P., Brunsteiner, M., Khinast, J., 2014. Investigation of migrant-

polymer interaction in pharmaceutical packaging material using the linear interaction energy algorithm. J Pharm Sci 103, 3197–3204. https://doi.org/10.1002/jps.24115

Feenstra, P., Brunsteiner, M., Khinast, J., 2012. Prediction of drug-packaging interactions via molecular dynamics (MD) simulations. Int J Pharm 431, 26–32. https://doi.org/10.1016/j.ijpharm.2012.03.049Francisco, M., Remolona, M., 2018. HOLMES: A Hybrid Ontology-Learning Materials Engineering System (Doctor of Philosophy). Columbia University, New York.

Gaddem, M.R., Kim, J., Matsunami, K., Hayashi, Y., Badr, S., Sugiyama, H., 2024. Roles of mechanistic, data-driven, and hybrid modeling approaches for pharmaceutical process design and operation. Curr Opin Chem Eng 44, 101019. https://doi.org/10.1016/j.coche.2024.101019

Garani, G., Chernov, A. V., Savvas, I.K., Butakova, M.A., 2019. A Data Warehouse Approach for Business Intelligence, in: Proceedings - 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2019. Institute of Electrical and Electronics Engineers Inc., pp. 70–75. https://doi.org/10.1109/WETICE.2019.00022Garcia, T., Cook, G., Nosal, R., 2008. PQLI key topics - Criticality, design space, and control strategy. J Pharm Innov. https://doi.org/10.1007/s12247-008-9032-4

Gasteiger, J., 2020. Chemistry in Times of Artificial Intelligence. ChemPhysChem. https://doi.org/10.1002/cphc.202000518

Ghobakhloo, M., Iranmanesh, M., Tseng, M.L., Grybauskas, A., Stefanini, A., Amran, A., 2023. Behind the definition of Industry 5.0: a systematic review of technologies, principles, components, and values. Journal of Industrial and Production Engineering 40, 432–447. https://doi.org/10.1080/21681015.2023.2216701

GlaxoSmithKline plc., 2022. GSK Onyx [WWW Document]. GSK. URL https://gskonyx.tech/ (accessed 8.15.24).

Golub, A., Lewis, J., Tinkoff, D., der Veken, L., Zhu, D., 2023. Rewired pharma companies will win in the digital age [WWW Document]. Mckinsey & Co. URL https://www.mckinsey.com/industries/life-sciences/our-insights/rewired-pharma-companies-will-win-in-the-digital-age (accessed 9.4.23).

Gregoire, J.M., Zhou, L., Haber, J.A., 2023. Combinatorial synthesis for AI-driven materials discovery. Nature Synthesis. https://doi.org/10.1038/s44160-023-00251-4

Gröls, J.R., Castro-Dominguez, B., 2021. Mechanochemical co-crystallization: Insights and predictions. Comput Chem Eng 153. https://doi.org/10.1016/j.compchemeng.2021.107416

Grom, M., Stavber, G., Drnovšek, P., Likozar, B., 2016. Modelling chemical kinetics of a complex reaction network of active pharmaceutical ingredient (API) synthesis with process optimization for benzazepine heterocyclic compound. Chemical Engineering Journal 283, 703–716. https://doi.org/10.1016/j.cej.2015.08.008

Guo, D., Srai, J.S., Settanni, E., 2024. Digital twin-enabled production-logistics synchronisation for enhancing efficiency and sustainability in pharmaceutical manufacturing., in: The 2024 International Conference on Automation in Manufacturing, Transportation and Logistics (ICaMaL). IEEE, Hong Kong, Hong Kong.

Hailemariam, L., Venkatasubramanian, V., 2010. Purdue ontology for pharmaceutical engineering: Part I. Conceptual framework. J Pharm Innov 5, 88–99. https://doi.org/10.1007/s12247-010-9081-3

Harmon, J.E., 2023. Argonne's self-driving lab accelerates the discovery process for materials with multiple applications [WWW Document]. Argonne National Laboratory. URL https://www.anl.gov/article/argonnes-selfdriving-lab-accelerates-the-discovery-process-for-materials-with-multiple-applications (accessed 7.29.24).Harrison, R.K., 2016. Phase II and phase III failures: 2013-2015. Nat Rev Drug Discov 15, 817–818. https://doi.org/10.1038/nrd.2016.184

Hatcher, L.E., Burgess, A.J., Payne, P., Wilson, C.C., 2020. From structure to crystallisation and pharmaceutical manufacturing: The CSD in CMAC workflows. CrystEngComm 22, 7475–7489. https://doi.org/10.1039/d0ce00898b

Hauck, T., 2024. 7 Digital Adoption Strategies That Boost Business Success [WWW Document]. Prosci. URL https://www.prosci.com/blog/digital-adoption-strategies (accessed 8.5.24).

Hausberg, J.P., Liere-Netheler, K., Packmohr, S., Pakura, S., Vogelsang, K., 2019. Research streams on digital transformation from a holistic business perspective: a systematic literature review and citation network analysis. Journal of Business Economics 89, 931–963. https://doi.org/10.1007/s11573-019-00956-z

Hechler, E., Weihrauch, M., Wu, Y., 2023. Data Fabric and Data Mesh Approaches with AI, Data Fabric and Data Mesh Approaches with AI. Apress. https://doi.org/10.1007/978-1-4842-9253-2Henderson, R.K., Jiménez-González, C., Constable, D.J.C., Alston, S.R., Inglis, G.G.A., Fisher, G., Sherwood, J., Binks, S.P., Curzons, A.D., 2011. Expanding GSK's solvent selection guide – embedding sustainability into solvent selection starting at medicinal chemistry. Green Chemistry 13, 854–862. https://doi.org/10.1039/c0gc00918kHlupić, T., Oreščanin, D., Ružak, D., Baranović, M., 2022. An Overview of Current Data Lake Architecture Models, in: MIPRO 2022. Opatija, Croatia, pp. 1082–1087.

Hole, G., Hole, A.S., McFalone-Shaw, I., 2021. Digitalization in pharmaceutical industry: What to focus on under the digital implementation process? Int J Pharm X. https://doi.org/10.1016/j.ijpx.2021.100095

Hu, H., Singh, A.N., Lehnherr, D., Mdluli, V., Chun, S.W., Makarewicz, A.M., Gouker, J.R., Ukaegbu, O., Li, S., Wen, X., McLaren, D.G., Velasquez, J.E., Moore, J.C., Galanie, S., Appiah-Amponsah, E., Regalado, E.L., 2024. Accelerating Pharmaceutical Process Development with an Acoustic Droplet Ejection-Multiple Reaction Monitoring-Mass Spectrometry Workflow. Anal Chem 96, 1138–1146. https://doi.org/10.1021/acs.analchem.3c04211

Hughes, L.D., Tsueng, G., DiGiovanna, J., Horvath, T.D., Rasmussen, L.V., Savidge, T.C., Stoeger, T., Turkarslan, S., Wu, Q., Wu, C., Su, A.I., Pache, L., The NIAID Systems Biology Data Dissemination Working Group, 2023. Addressing barriers in FAIR data practices for biomedical data. Sci Data 10, 98. https://doi.org/10.1038/s41597-023-01969-8Humayed, A., Lin, J., Li, F., Luo, B., 2017. Cyber-Physical Systems Security - A Survey. IEEE Internet Things J 4, 1802–1831. https://doi.org/10.1109/JIOT.2017.2703172

IBM, 2024. What is a data fabric? [WWW Document]. IBM. URL https://www.ibm.com/topics/data-fabric (accessed 7.30.24).

ICH Secretariat with the ICH Management Committee, MedDRA Management Committee, 2021. ICH ASSOCIATION 2020 ANNUAL REPORT.

Içten, E., Maloney, A.J., Beaver, M.G., Shen, D.E., Zhu, X., Graham, L.R., Robinson, J.A., Huggins, S., Allian, A., Hart, R., Walker, S.D., Rolandi, P., Braatz, R.D., 2020. A Virtual Plant for Integrated Continuous Manufacturing of a Carfilzomib Drug Substance Intermediate, Part 1: CDI-Promoted Amide Bond Formation. Org Process Res Dev 24, 1861–1875. https://doi.org/10.1021/acs.oprd.0c00187

International Council for Harmonisation, 2020. ICH Press Release: ICH Global Harmonisation Efforts Continued Through Virtual Meetings. International Council for Harmonisation .

International Council for Harmonisation, 2019. ICH Harmonised Guideline: Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management Q12, in: International Conference on Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.

International Council for Harmonisation, 2012. ICH Harmonised Tripartite Guideline: Development and Manufacture of Drug Substances (Chemical Entities and Biotechnological/Biological Entities) Q11, in: International Conference on Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.

International Council for Harmonisation, 2009. ICH Harmonised Tripartite Guideline: Pharmaceutical Development Q8(R2), in: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.

International Council for Harmonisation., 2008. ICH Harmonised Tripartite Guideline : Pharmaceutical Quality System Q10, in: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.

International Council for Harmonisation, 2005. ICH Harmonised Tripartite Guideline: Quality Risk Management Q9, in: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, 2023. ICH Harmonised Guideline: Analytical Procedure Development Q14. International Conference for Harmonisation.

Intrepid, 2024. Combining artificial intelligence and robotics to develop next-generation medicines [WWW Document]. Intrepid. URL https://intrepidlabs.tech/platform (accessed 7.29.24).

Ismail, H.Y., Singh, M., Darwish, S., Kuhs, M., Shirazian, S., Croker, D.M., Khraisheh, M., Albadarin, A.B., Walker, G.M., 2019. Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time distribution in twin-screw wet granulation. Powder Technol 343, 568–577. https://doi.org/10.1016/j.powtec.2018.11.060

Jafari, N., Azarian, M., Yu, H., 2022. Moving from Industry 4.0 to Industry 5.0: What Are the Implications for Smart Logistics? Logistics. https://doi.org/10.3390/logistics6020026

Jamei, M., Marciniak, S., Feng, K., Barnett, A., Tucker, G., Rostami-Hodjegan, A., 2009. The Simcyp® Population-based ADME Simulator. Expert Opin Drug Metab Toxicol 5, 211–233. https://doi.org/https://doi.org/10.1517/17425250802691074

Jang, E.H., Park, Y.S., Kim, M.S., Choi, D.H., 2020. Model-based scale-up methodologies for pharmaceutical granulation. Pharmaceutics. https://doi.org/10.3390/pharmaceutics12050453

Jayatunga, M.K.P., Xie, W., Ruder, L., Schulze, U., Meier, C., 2022. AI in small-molecule drug discovery: a coming wave? Nat Rev Drug Discov. https://doi.org/10.1038/d41573-022-00025-1

Jiménez-Luna, J., Grisoni, F., Weskamp, N., Schneider, G., 2021. Artificial intelligence in drug discovery: recent advances and future perspectives. Expert Opin Drug Discov 16, 949–959. https://doi.org/10.1080/17460441.2021.1909567

Jing, Y., Bian, Y., Hu, Z., Wang, L., Xie, X.Q.S., 2018. Deep Learning for Drug Design: an Artificial Intelligence Paradigm for Drug Discovery in the Big Data Era. AAPS Journal. https://doi.org/10.1208/s12248-018-0210-0

Jones, N., 2024. A Vision of the Future: Machine Learning in Packaging Inspection [WWW Document]. Packaging Digest. URL https://www.packagingdigest.com/automation/a-vision-of-the-future-machine-learning-in-packaging-inspection (accessed 9.23.24).

Juran, J.M., 1992. Juran on Quality by Design: The New Steps for Planning Quality into Goods and Services. Free Press, Maxwell Macmillan Canada, Maxwell Macmillan International, New York, Toronto, New York.

Kar, M., Chourasiya, Y., Maheshwari, R., Tekade, R.K., 2018. Current developments in excipient science: Implication of quantitative selection of each excipient in product development, in: Basic Fundamentals of Drug Delivery. Elsevier, pp. 29–83. https://doi.org/10.1016/B978-0-12-817909-3.00002-9

Kavasidis, I., Lallas, E., Leligkou, H.C., Oikonomidis, G., Karydas, D., Gerogiannis, V.C., Karageorgos, A., 2023. Deep Transformers for Computing and Predicting ALCOA+Data Integrity Compliance in the Pharmaceutical Industry. Applied Sciences (Switzerland) 13. https://doi.org/10.3390/app13137616Kitsios, F., Kamariotou, M., 2021. Artificial intelligence and business strategy towards digital transformation: A research agenda. Sustainability (Switzerland) 13, 1–16. https://doi.org/10.3390/su13042025

Kodam, M., Curtis, J., Hancock, B., Wassgren, C., 2012. Discrete element method modeling of bi-convex pharmaceutical tablets: Contact detection algorithms and validation. Chem Eng Sci 69, 587–601. https://doi.org/10.1016/j.ces.2011.11.011

Lakerveld, R., Benyahia, B., Heider, P.L., Zhang, H., Wolfe, A., Testa, C.J., Ogden, S., Hersey, D.R., Mascia, S., Evans, J.M.B., Braatz, R.D., Barton, P.I., 2015. The Application of an Automated Control Strategy for an Integrated Continuous Pharmaceutical Pilot Plant. Org Process Res Dev 19, 1088–1100. https://doi.org/10.1021/op500104d

Lavecchia, A., 2015. Machine-learning approaches in drug discovery: Methods and applications. Drug Discov Today. https://doi.org/10.1016/j.drudis.2014.10.012

Leane, M., Pitt, K., Reynolds, G., Anwar, J., Charlton, S., Crean, A., Creekmore, R., Davies, C., DeBeer, T., De-Matas, M., Djemai, A., Douroumis, D., Gaisford, S., Gamble, J., Stone, E.H., Kavanagh, A., Khimyak, Y., Kleinebudde, P., Moreton, C., Paudel, A., Storey, R., Toschkoff, G., Vyas, K., 2015. A proposal for a drug product Manufacturing Classification System (MCS) for oral solid dosage forms. Pharm Dev Technol. https://doi.org/10.3109/10837450.2014.954728

Leane, M., Pitt, K., Reynolds, G., Tantuccio, A., Moreton, C., Crean, A., Kleinebudde, P., Carlin, B., Gamble, J., Gamlen, M., Stone, E., Kuentz, M., Gururajan, B., Khimyak, Y.Z., Van Snick, B., Andersen, S., Misic, Z., Peter, S., Sheehan, S., 2024. Ten years of the manufacturing classification system: a review of literature applications and an extension of the framework to

continuous manufacture. Pharm Dev Technol. https://doi.org/10.1080/10837450.2024.2342953

Leane, M., Pitt, K., Reynolds, G.K., Dawson, N., Ziegler, I., Szepes, A., Crean, A.M., Dall Agnol, R., The Manufacturing Classification Sy, 2018. Manufacturing classification system in the real world: factors influencing manufacturing process choices for filed commercial oral solid dosage formulations, case studies from industry and considerations for continuous processing. Pharm Dev Technol 23, 964–977. https://doi.org/10.1080/10837450.2018.1534863

Lee, E.A., 2006. Cyber-Physical Systems-Are Computing Foundations Adequate?, in: NSF Workshop On Cyber-Physical Systems: Research Motivation, Techniques and Roadmap.

Li, H., Alkahtani, M.E., Basit, A.W., Elbadawi, M., Gaisford, S., 2023. Optimizing environmental sustainability in pharmaceutical 3D printing through machine learning. Int J Pharm 648. https://doi.org/10.1016/j.ijpharm.2023.123561

Litster, J., Bogle, I.D.L., 2019. Smart Process Manufacturing for Formulated Products. Engineering. https://doi.org/10.1016/j.eng.2019.02.014

Liu, J., Benyahia, B., 2021. Systematic model-based dynamic optimization of a combined cooling and antisolvent multistage continuous crystallization process, in: Computer Aided Chemical Engineering. Elsevier B.V., pp. 1221–1227. https://doi.org/10.1016/B978-0-323-88506-5.50188-1

Liu, M., Fang, S., Dong, H., Xu, C., 2021. Review of digital twin about concepts, technologies, and industrial applications. J Manuf Syst 58, 346–361. https://doi.org/10.1016/j.jmsy.2020.06.017

Liu, Q., Huang, R., Hsieh, J., Zhu, H., Tiwari, M., Liu, G., Jean, D., ElZarrad, M.K., Fakhouri, T., Berman, S., Dunn, B., Diamond, M.C., Huang, S.M., 2023. Landscape Analysis of the Application of Artificial Intelligence and Machine Learning in Regulatory Submissions for Drug Development From 2016 to 2021. Clin Pharmacol Ther. https://doi.org/10.1002/cpt.2668

Lo, Y.C., Rensi, S.E., Torng, W., Altman, R.B., 2018. Machine learning in chemoinformatics and drug discovery. Drug Discov Today. https://doi.org/10.1016/j.drudis.2018.05.010Lomax, J., 2022. What Are Ontologies and How Are They Creating a FAIRer Future for the Life Sciences? [WWW Document]. Technology Networks Informatics. URL https://www.technologynetworks.com/informatics/articles/what-are-ontologies-and-how-are-they-creating-a-fairer-future-for-the-life-sciences-363371 (accessed 3.26.24).

Lorenz, S., Amsel, A.K., Puhlmann, N., Reich, M., Olsson, O., Kümmerer, K., 2021. Toward Application and Implementation of in Silico Tools and Workflows within Benign by Design Approaches. ACS Sustain Chem Eng. https://doi.org/10.1021/acssuschemeng.1c03070Loschen, C., Klamt, A., 2015. Solubility prediction, solvate and cocrystal screening as tools for rational crystal engineering. Journal of Pharmacy and Pharmacology 67, 803–811. https://doi.org/10.1111/jphp.12376

Ma, C.Y., Roberts, K.J., 2019. Morphological population balance modelling of the effect of crystallisation environment on the evolution of crystal size and shape of para-aminobenzoic acid. Comput Chem Eng 126, 356–370. https://doi.org/10.1016/j.compchemeng.2019.04.019

Maclean, N., Armstrong, J.A., Carroll, M.A., Salehian, M., Mann, J., Reynolds, G., Johnston, B., Markl, D., 2024. Flexible modelling of the dissolution performance of directly compressed tablets. Int J Pharm 656, 124084. https://doi.org/10.15129/0de306

Madarász, L., Mészáros, L.A., Köte, Á., Farkas, A., Nagy, Z.K., 2023. AI-based analysis of in-line process endoscope images for real-time particle size measurement in a continuous pharmaceutical milling process. Int J Pharm 641. https://doi.org/10.1016/j.ijpharm.2023.123060

Maddikunta, P.K.R., Pham, Q.V., B, P., Deepa, N., Dev, K., Gadekallu, T.R., Ruby, R., Liyanage, M., 2022. Industry 5.0: A survey on enabling technologies and potential applications. J Ind Inf Integr. https://doi.org/10.1016/j.jii.2021.100257

Mahjour, B., Hoffstadt, J., Cernak, T., 2023. Designing Chemical Reaction Arrays Using Phactor and ChatGPT. Org Process Res Dev 27, 1510–1516. https://doi.org/10.1021/acs.oprd.3c00186

Marmolejo-Saucedo, J.A., 2020. Design and Development of Digital Twins: a Case Study in Supply Chains. Mobile Networks and Applications 25, 2141–2160. https://doi.org/10.1007/s11036-020-01557-9

Martin, N.L., Schomberg, A.K., Finke, J.H., Abraham, T.G.M., Kwade, A., Herrmann, C., 2021. Process modeling and simulation of tableting—an agent-based simulation methodology for direct compression. Pharmaceutics 13. https://doi.org/10.3390/pharmaceutics13070996

Marwedel, P., 2021. Embedded System Design Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things, 4th ed.

Matić, J., Stanković-Brandl, M., Bauer, H., Lovey, J., Martel, S., Herkenne, C., Paudel, A., Khinast, J., 2023. Pharmaceutical hot melt extrusion process development using QbD and digital twins. Int J Pharm 631. https://doi.org/10.1016/j.ijpharm.2022.122469

Medicines & Healthcare products Regulatory Agency, 2024. Impact of AI on the regulation of medical products: Implementing the AI White Paper principles.

Medicines & Healthcare products Regulatory Agency, 2023. Guidance: Software and AI as a Medical Device Change Programme - Roadmap [WWW Document]. UK Government. URL https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap (accessed 9.9.24).

Medicines & Healthcare products Regulatory Agency, Brunel University, 2022. Research into methodology for determining significant change in the way that an adaptive AI algorithm medical device is working and how such change should be regulated.Medina-Franco, J.L., 2021. Grand Challenges of Computer-Aided Drug Design: The Road Ahead. Frontiers in Drug Discovery 1. https://doi.org/10.3389/fddsv.2021.728551

Mollick, E., 2022. ChatGPT Is a Tipping Point for AI [WWW Document]. Harv Bus Rev. URL https://hbr.org/2022/12/chatgpt-is-a-tipping-point-for-ai (accessed 5.16.23).

Moreno-Benito, M., Lee, K.T., Kaydanov, D., Verrier, H.M., Blackwood, D.O., Doshi, P., 2022. Digital twin of a continuous direct compression line for drug product and process design using a hybrid flowsheet modelling approach. Int J Pharm 628. https://doi.org/10.1016/j.ijpharm.2022.122336

Mswahili, M.E., Lee, M.J., Martin, G.L., Kim, J., Kim, P., Choi, G.J., Jeong, Y.S., 2021. Cocrystal prediction using machine learning models and descriptors. Applied Sciences (Switzerland) 11, 1–12. https://doi.org/10.3390/app11031323

Nagaprasad, S., Padmaja, D.L., Qureshi, Y., Bangare, S.L., Mishra, M., Mazumdar, B.D., 2021. Investigating the Impact of Machine Learning in Pharmaceutical Industry. J Pharm Res Int 6–14. https://doi.org/10.9734/jpri/2021/v33i46a32834

Nahavandi, S., 2019. Industry 5.0-a human-centric solution. Sustainability (Switzerland) 11. https://doi.org/10.3390/su11164371Nair, A., Loveday, K.A., Kenyon, C., Qu, J., Kis, Z., 2024. Quality by Digital Design for Developing Platform RNA Vaccine and Therapeutic Manufacturing Processes, in: Kramps, T. (Ed.), RNA Vaccines. Humana Press, New York, NY, USA, pp. 339–364. https://doi.org/https://doi.org/10.1007/978-1-0716-3770-8Nakapraves, S., Warzecha, M., Mustoe, C.L., Srirambhatla, V., Florence, A.J., 2022. Prediction of Mefenamic Acid Crystal Shape by Random Forest Classification. Pharm Res 39, 3099–3111. https://doi.org/10.1007/s11095-022-03450-4

Nambiar, A., Mundra, D., 2022. An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management. Big Data and Cognitive Computing. https://doi.org/10.3390/bdcc6040132Naveršnik, K., Bohanec, S., 2008. Predicting drug hydrolysis based on moisture uptake in various packaging designs. European Journal of Pharmaceutical Sciences 35, 447–456. https://doi.org/10.1016/j.ejps.2008.09.007

Niederer, S.A., Sacks, M.S., Girolami, M., Willcox, K., 2021. Scaling digital twins from the artisanal to the industrial. Nat Comput Sci. https://doi.org/10.1038/s43588-021-00072-5

Nielsen, R.F., Nazemzadeh, N., Sillesen, L.W., Andersson, M.P., Gernaey, K. V., Mansouri, S.S., 2020. Hybrid machine learning assisted modelling framework for particle processes. Comput Chem Eng 140. https://doi.org/10.1016/j.compchemeng.2020.106916

Onaji, I., Tiwari, D., Soulatiantork, P., Song, B., Tiwari, A., 2022. Digital twin in manufacturing: conceptual framework and case studies. Int J Comput Integr Manuf. https://doi.org/10.1080/0951192X.2022.2027014

Ong, J.J., Castro, B.M., Gaisford, S., Cabalar, P., Basit, A.W., Pérez, G., Goyanes, A., 2022. Accelerating 3D printing of pharmaceutical products using machine learning. Int J Pharm X 4. https://doi.org/10.1016/j.ijpx.2022.100120

Open AI, 2023. Introducing ChatGPT. [WWW Document]. URL https://openai.com/blog/chatgpt (accessed 5.16.23).

Open Source Robotics Foundation, 2024. Robotic Operating Systems [WWW Document]. Robotic Operating Systems. URL https://roscon.ros.org/2024/ (accessed 10.8.24).

Ottoboni, S., Brown, C.J., Mehta, B., Jimeno, G., Mitchell, N.A., Sefcik, J., Price, C.J., 2022. Digital Design of Filtration and Washing of Active Pharmaceutical Ingredients via Mechanistic Modeling. Org Process Res Dev 26, 3236–3253. https://doi.org/10.1021/acs.oprd.2c00165

Ottoboni, S., Wareham, B., Vassileiou, A., Robertson, M., Brown, C.J., Johnston, B., Price, C.J., 2021. A Novel Integrated Workflow for Isolation Solvent Selection Using Prediction and Modeling. Org Process Res Dev 25, 1143–1159. https://doi.org/10.1021/acs.oprd.0c00532Oyebola, A., Opeoluwa, A.B., 2015. An Ontological Informatics on Development of Tablet Drug in Pharmaceutical Industry. International Journal of Advanced Research in Computer Science 6.

Parrott, N., Lavé, T., 2002. Prediction of intestinal absorption: comparative assessment of gastroplusTM and ideaTM. European Journal of Pharmaceutical Sciences 17, 51–61. https://doi.org/https://doi.org/10.1016/S0928-0987(02)00132-X

Patel, L., Shukla, T., Huang, X., Ussery, D.W., Wang, S., 2020. Machine Learning Methods in Drug Discovery. Molecules 25. https://doi.org/10.3390/MOLECULES25225277

Patel, V., Shah, M., 2022. Artificial intelligence and machine learning in drug discovery and development. Intelligent Medicine. https://doi.org/10.1016/j.imed.2021.10.001

Pauli, T., Williams, R., 2018. Future Pharma 2018: Digitally transforming pharma's regulatory processes [WWW Document]. URL https://deep-dive.pharmaphorum.com/magazine/future-pharma-2018/digitally-transforming-pharmas-regulatory-processes/ (accessed 3.4.24).

Pereira Diaz, L., Brown, C.J., Ojo, E., Mustoe, C., Florence, A.J., 2023. Machine learning approaches to the prediction of powder flow behaviour of pharmaceutical materials from physical properties. Digital Discovery 2, 692–701. https://doi.org/10.1039/D2DD00106CPickles, T., Mustoe, C., Boyle, C., Cardona, J., Brown, C.J., Florence, A.J., 2024. Developing a model-driven workflow for the digital design of small-scale batch cooling crystallisation with the antiviral lamivudine. CrystEngComm 26, 822–834. https://doi.org/10.1039/d3ce00897e

Pickles, T., Mustoe, C., Brown, C., Florence, A., 2022a. Developing an autonomous DataFactory workflow for small-scale batch cooling crystallisation with the antiviral lamivudine, in: CMAC Open Day 2022. Glasgow.

Pickles, T., Mustoe, C., Brown, C., Florence, A., 2022b. Autonomous DataFactory: High-throughput screening for large-scale data collection to inform medicine manufacture. British Journal of Pharmacy 7. https://doi.org/10.5920/bjpharm.1128

Popov, V. V., Kudryavtseva, E. V., Katiyar, N.K., Shishkin, A., Stepanov, S.I., Goel, S., 2022. Industry 4.0 and Digitalisation in Healthcare. Materials. https://doi.org/10.3390/ma15062140Potter, C.J., 2022. European Medicines Agency Approach to Facilitating Innovative Manufacturing Approaches [WWW Document]. ISPE International Society for Pharmaceutical Engineering. URL https://ispe.org/pharmaceutical-engineering/ispeak/european-medicines-agency-approach-facilitating-innovative (accessed 2.20.23).

Powder Systems, 2024. Ensuring quality and reliability: IQ, OQ, and PQ protocols in pharmaceutical manufacturing [WWW Document]. News Medical & Life Sciences.

Prasad, E., Robertson, J., Halbert, G.W., 2022. Mefenamic acid solid dispersions: Impact of formulation composition on processing parameters, product properties and performance. Int J Pharm 616. https://doi.org/10.1016/j.ijpharm.2022.121505Precision Solutions Inc., 2024. IQ OQ PQ for Pharmaceutical [WWW Document]. Precision Solutions Inc. URL https://www.precisionsolutionsinc.com/iq-oq-pq-for-pharmaceutical/ (accessed 5.30.24).

Ramos, M.C., Collison, C.J., White, A.D., 2024. A Review of Large Language Models and Autonomous Agents in Chemistry.Rane, N., 2023. Transformers in Industry 4.0, Industry 5.0, and Society 5.0: Roles and Challenges. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4609915

Ranjan Yadav, P., Iqbal Nasiri, M., Vora, L.K., Larrañeta, E., Donnelly, R.F., Pattanayek, S.K., Bhusan Das, D., 2022. Super-swelling hydrogel-forming microneedle based transdermal

drug delivery: Mathematical modelling, simulation and experimental validation. Int J Pharm 622. https://doi.org/10.1016/j.ijpharm.2022.121835

Ravendra Singh, 2024. Pharma 4.0: Advanced Continuous Pharmaceutical Tablet Manufacturing [WWW Document]. Pharma Focus Europe. URL https://www.pharmafocuseurope.com/manufacturing/pharma-advanced-continuous-pharmaceutical (accessed 4.3.24).

Redrup Hill, E., Mitchell, C., Myles, P., Branson, R., Frangi, A.F., 2023. Cross-Regulator Workshop: Journeys, experiences and best practices on computer modelled and simulated regulatory evidence— Workshop Report. InSilicoUK Pro-Innovation Regulations Network. https://doi.org/10.5281/zenodo.10121103

Rege, P.D., Schuster, A., Lamerz, J., Moessner, C., Göhring, W., Hidber, P., Stahr, H., Andrei, O.M., Burren, J., Moesching, A., Coleman, D., Hildbrand, S., 2024. QbD Approach to Process Characterization and Quantitative Criticality Assessment of Process Parameters. Org Process Res Dev. https://doi.org/10.1021/acs.oprd.3c00356Remmelgas, J., 2017. Predicting Moisture Uptake in Solid Dosage Packaging. Pharmaceutical Technology 41, 44–49.

Riaz Ahmed, K., Pathmanathan, P., Kabadi, S. V., Van Doren, J., Kruhlak, N., Lumen, A., Martinez, M., Morrison, T., Schuette, P., Tegenge, M., Aungst, J., Bandele, O., Chen, Y., Craven, B., Dennis, S., Ditto, M., Ellis, C., Fanaselle, W., Fisher, J., Hariharan, P., Hong, H., Kainz, W., Kuemmel, C., Liu, Y., Mattia, A., Matthews, E., Oryang, D., Pang, H., Paredes, A., Rodriguez, P., Sakkiah, S., Tian, G., Verma, R., Volarath, P., Xu, X., Yang, H., 2022. Successes and Opportunities in Modeling & Simulation for FDA. Modeling & Simulation Working Group of the Senior Science Council, U.S. Food and Drug Administration.

Robotics Growth Partnership, 2022. The Cyber-Physical Infrastructure.

Rogers, L., Briggs, N., Achermann, R., Adamo, A., Azad, M., Brancazio, D., Capellades, G., Hammersmith, G., Hart, T., Imbrogno, J., Kelly, L.P., Liang, G., Neurohr, C., Rapp, K., Russell, M.G., Salz, C., Thomas, D.A., Weimann, L., Jamison, T.F., Myerson, A.S., Jensen, K.F., 2020. Continuous Production of Five Active Pharmaceutical Ingredients in Flexible Plug-and-Play Modules: A Demonstration Campaign. Org Process Res Dev 24, 2183–2196. https://doi.org/10.1021/acs.oprd.0c00208

Romañach, R.J., Stelzer, T., Sanchez, E., Muzzio, F., 2023. Advanced pharmaceutical manufacturing: A functional definition. J Adv Manuf Process 5. https://doi.org/10.1002/amp2.10150

Ruan, Y., Lu, C., Xu, N., Zhang, J., Xuan, J., Pan, J., Fang, Q., Gao, H., Shen, X., Ye, N., Zhang, Q., Mo, Y., 2024. Accelerated end-to-end chemical synthesis development with large language models. Chemrxiv. https://doi.org/10.26434/chemrxiv-2024-6wmg4

Sadeghi, A., Su, C.H., Khan, A., Lutfor Rahman, M., Sani Sarjadi, M., Sarkar, S.M., 2022. Machine learning simulation of pharmaceutical solubility in supercritical carbon dioxide: Prediction and experimental validation for busulfan drug. Arabian Journal of Chemistry 15. https://doi.org/10.1016/j.arabjc.2021.103502

Samson, Y., 2021. ALCOA+ - what does it mean? [WWW Document]. ECA Foundation. URL https://www.gmp-compliance.org/gmp-news/alcoa-what-does-it-mean (accessed 8.29.24).

Sanislav, T., Miclea, L., 2012. Cyber-Physical Systems-Concept, Challenges and Research Areas. Journal of Control Engineering and Applied Informatics 14, 28–33.

Sansare, S., Duran, T., Mohammadiarani, H., Goyal, M., Yenduri, G., Costa, A., Xu, X., O'Connor, T., Burgess, D., Chaudhuri, B., 2021. Artificial neural networks in tandem with molecular descriptors as predictive tools for continuous liposome manufacturing. Int J Pharm 603. https://doi.org/10.1016/j.ijpharm.2021.120713

Santos, J.A.M., Lopes, M.R., Viegas, J.L., Vieira, S.M., Sousa, J.M.C., 2020. Internal supply chain digital twin of a pharmaceutical company, in: IFAC-PapersOnLine. Elsevier B.V., pp. 10797–10802. https://doi.org/10.1016/j.ifacol.2020.12.2864

Schlander, M., Hernandez-Villafuerte, K., Cheng, C.Y., Mestre-Ferrandiz, J., Baumann, M., 2021. How Much Does It Cost to Research and Develop a New Drug? A Systematic Review and Assessment. Pharmacoeconomics. https://doi.org/10.1007/s40273-021-01065-y

Schrimpf, G., 2022. Gerresheimer and Merck Transform Primary Packaging into Digital Twins. Merck KGaA.

Schroeder, S.L.M., 2024. Homogeneous Organic Crystal Nucleation Rates in Solution from the Perspective of Chemical Reaction Kinetics. Crystals (Basel) 14. https://doi.org/10.3390/cryst14040349

Seenivasan Mphasis, D., Seenivasan, D., 2023. Real-Time Data Processing with Streaming ETL. Article in International Journal of Science and Research. https://doi.org/10.21275/SR24619000026Sembiring, M.H., Novagusda, F.N., 2024. Enhancing Data Security Resilience in AI-Driven Digital Transformation: Exploring Industry Challenges and Solutions Through ALCOA+ Principles. Acta Informatica Medica 32, 65–70. https://doi.org/10.5455/aim.2024.32.65-70

Siemens, 2025a. gPROMS FormulatedProducts [WWW Document]. Siemens. URL https://www.siemens.com/global/en/products/automation/industry-software/gproms-digital-process-design-and-operations/gproms-modelling-environments/gproms-formulatedproducts.html (accessed 3.17.25).Siemens, 2025b. gPROMS – Digital Process Twin technology [WWW Document]. Siemens. URL https://www.siemens.com/global/en/products/automation/industry-software/gproms-digital-process-design-and-operations.html (accessed 3.17.25).SiLA, 2018. The SiLA Consortium [WWW Document]. SiLA. URL https://sila-standard.com/about-us/ (accessed 10.8.24).

Silva, M.I., Khadra, I., Pyper, K., Halbert, G.W., 2023. Fed intestinal solubility limits and distributions applied to the Developability classification system. European Journal of Pharmaceutics and Biopharmaceutics 186, 74–84. https://doi.org/10.1016/j.ejpb.2023.03.005

Simon Hart, 2023. Cyber-physical infrastructure: what leaders need to know [WWW Document]. UKRI. URL https://www.ukri.org/blog/cyber-physical-infrastructure-what-leaders-need-to-know/ (accessed 10.10.24).

SimulationsPlus, 2023. GastroPlus® PBPK & PBBM Modeling and Simulation [WWW Document]. GastroPlus®. URL https://www.simulations-plus.com/software/gastroplus/ (accessed 9.19.23).

Singhal, B., Aggarwal, A., 2022. ETL, ELT and Reverse ETL: A business case Study, in: 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE). IEEE, pp. 1–4. https://doi.org/10.1109/ICATIECE56365.2022.10046997

Sperry, J.B., Azuma, M., Stone, S., 2021. Explosive Hazard Identification in Pharmaceutical Process Development: A Novel Screening Method and Workflow for Shipping Potentially Explosive Materials. Org Process Res Dev 25, 212–224. https://doi.org/10.1021/acs.oprd.0c00467

Srai, J.S., Bauer, P., Badman, C., Bresciani, M., Cooney, C.L., Florence, A., Hausner, D., Konstantinov, K., Lee, S.L., Mascia, S., Nasr, M., Trout, B.L., 2024. Emerging Applications and Regulatory Strategies for Advanced Medicines Manufacturing – Towards the Development of a Platform Approach. J Pharm Sci. https://doi.org/10.1016/j.xphs.2024.04.016

Srai, J.S., Settanni, E., Aulakh, P.K., 2020. Evaluating the Business Case for Continuous Manufacturing of Pharmaceuticals: A Supply Network Perspective., in: Nagy, Z., El Hagrasy, A., Litster, J. (Eds.), Continuous Pharmaceutical Processing. AAPS Advances in the Pharmaceutical Sciences Series. American Association of Pharmaceutical Scientists, pp. 477–512.

Srai, J.S., Settanni, E., Tsolakis, N., Aulakh, K., 2019. Supply Chain Digital Twins: Opportunities and Challenges Beyond the Hype 26–27.

Strömert, P., Hunold, J., Castro, A., Neumann, S., Koepler, O., 2022. Ontologies4Chem: the landscape of ontologies in chemistry. Pure and Applied Chemistry 94, 605–622. https://doi.org/10.1515/pac-2021-2007Su, Q., Ganesh, S., Moreno, M., Bommireddy, Y., Gonzalez, M., Reklaitis, G. V., Nagy, Z.K., 2019. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. Comput Chem Eng 125, 216–231. https://doi.org/10.1016/j.compchemeng.2019.03.001

Subramanian, K., 2020. Digital Twin for Drug Discovery and Development—The Virtual Liver. J Indian Inst Sci. https://doi.org/10.1007/s41745-020-00185-2

Sun, D., Gao, W., Hu, H., Zhou, S., 2022. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. https://doi.org/10.1016/j.apsb.2022.02.002

Szilágyi, B., Eren, A., Quon, J.L., Papageorgiou, C.D., Nagy, Z.K., 2022. Digital Design of the Crystallization of an Active Pharmaceutical Ingredient Using a Population Balance Model with a Novel Size Dependent Growth Rate Expression. From Development of a Digital Twin to In Silico Optimization and Experimental Validation. Cryst Growth Des 22, 497–512. https://doi.org/10.1021/acs.cgd.1c01108

Tartal, J., U.S. Food & Drug Administration, 2015. Quality System Regulation Process Validation, in: Small Business Regulatory Education for Industry (REdI). Silver Spring, MD, USA.

Thantilage, R.D., Le-Khac, N.A., Kechadi, M.T., 2023. Healthcare data security and privacy in Data Warehouse architectures. Inform Med Unlocked. https://doi.org/10.1016/j.imu.2023.101270

The American Society of Mechanical Engineers, 2024. Verification, Validation and Uncertainty Quantification (VVUQ) [WWW Document]. The American Society of Mechanical Engineers. URL https://www.asme.org/codes-standards/publications-information/verification-validation-uncertainty (accessed 5.30.24).

The FDA Group, 2023. Computer System Validation (CSV) in the FDA-Regulated Industries [WWW Document]. The FDA Group. URL https://www.thefdagroup.com/blog/computer-system-validation (accessed 10.14.24).

The FDA Group LLC, 2024. A Basic Guide to IQ, OQ, PQ in FDA-Regulated Industries [WWW Document]. The FDA Group LLC.

The International Pharmaceutical Excipients Council (IPEC Federation), 2023. The International Pharmaceutical Excipients Council General GLOSSARY of Terms and Acronyms For Pharmaceutical Excipients. The International Pharmaceutical Excipients Council – Federation (IPEC Federation) asbl.

Tsopanoglou, A., Jiménez del Val, I., 2021. Moving towards an era of hybrid modelling: advantages and challenges of coupling mechanistic and data-driven models for upstream pharmaceutical bioprocesses. Curr Opin Chem Eng. https://doi.org/10.1016/j.coche.2021.100691Tubettificio Perfektup, 2023. Pharmaceutical packaging and artificial intelligence: how AI will revolutionize the industry [WWW Document]. Tubettificio Perfektup. URL https://www.perfektup.it/en/pharmaceutical-packaging-and-artificial-intelligence-how-ai-will-revolutionize-the-industry/ (accessed 9.23.24).

Tukamuhabwa, B.R., Stevenson, M., Busby, J., Zorzini, M., 2015. Supply chain resilience: Definition, review and theoretical foundations for further study. Int J Prod Res. https://doi.org/10.1080/00207543.2015.1037934

Tyagi, A.K., Sreenath, N., 2021. Cyber Physical Systems: Analyses, challenges and possible solutions. Internet of Things and Cyber-Physical Systems 1, 22–33. https://doi.org/10.1016/j.iotcps.2021.12.002

University of Liverpool, 2024. Materials Innovation Factory [WWW Document]. University of Liverpool. URL https://www.liverpool.ac.uk/materials-innovation-factory/research/ (accessed 9.12.24).University of York, 2024. Data architecture: the digital backbone of a business [WWW Document]. University of York. URL https://online.york.ac.uk/data-architecture-the-digital-backbone-of-a-business/ (accessed 7.27.24).

Unnikrishnan, S., Donovan, J., Macpherson, R., Tormey, D., 2021. In-process analysis of pharmaceutical emulsions using computer vision and artificial intelligence. Chemical Engineering Research and Design 166, 281–294. https://doi.org/10.1016/j.cherd.2020.12.010

Urwin, S.J., Chong, M.W.S., Li, W., McGinty, J., Mehta, B., Ottoboni, S., Pathan, M., Prasad, E., Robertson, M., McGowan, M., al-Attili, M., Gramadnikova, E., Siddique, M., Houson, I., Feilden, H., Benyahia, B., Brown, C.J., Halbert, G.W., Johnston, B., Nordon, A., Price, C.J., Reilly, C.D., Sefcik, J., Florence, A.J., 2023. Digital process design to define and deliver pharmaceutical particle attributes. Chemical Engineering Research and Design 196, 726–749. https://doi.org/10.1016/j.cherd.2023.07.003Urwin, S.J., Levilain, G., Marziano, I., Merritt, J.M., Houson, I., Ter Horst, J.H., 2020. A Structured Approach to Cope with Impurities during Industrial Crystallization Development. Org Process Res Dev 24, 1443–1456. https://doi.org/10.1021/acs.oprd.0c00166

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), 2023. ICH: Q13 Continuous Manufacturing of Drug Substances and Drug Products Guidance for Industry.

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), 2012. Q8, Q9, & Q10 Questions and Answers: Appendix: Q&As from Training Sessions.U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and

Research (CBER), Center for Veterinary Medicine (CVM), 2016. Data Integrity and Compliance With CGMP Guidance for Industry. Pharmaceutical Quality/Manufacturing Standards (CGMP).

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Veterinary Medicine (CVM), 2011. Guidance for Industry Process Validation: General Principles and Practices.

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), Office of Regulatory Affairs (ORA), 2004. Guidance for Industry PAT - A Framework for Innovative Pharmaceutical Development, manufacturing, and Quality Assurance [WWW Document]. URL https://www.fda.gov/media/71012/download (accessed 9.6.23).

U.S. Food & Drug Administration, 2024a. Using Artificial Intelligence & Machine Learning in the Development of Drug and Biological Products. U.S. Food & Drug Administration.

U.S. Food & Drug Administration, 2024b. Digital Health Center of Excellence [WWW Document]. U.S. Food & Drug Administration. URL https://www.fda.gov/medical-devices/digital-health-center-excellence (accessed 10.8.24).U.S. Food & Drug Administration, 2024c. Real-World Evidence [WWW Document]. U.S. Food & Drug Administration. URL https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence (accessed 10.8.24).

U.S. Food & Drug Administration, 2024d. Digital Health Technologies (DHTs) for Drug Development [WWW Document]. U.S. Food & Drug Administration. URL https://www.fda.gov/science-research/science-and-research-special-topics/digital-health-technologies-dhts-drug-development (accessed 10.8.24).

U.S. Food & Drug Administration, 2024e. Artificial Intelligence and Machine Learning (AI/ML) for Drug Development [WWW Document]. U.S. Food & Drug Administration. URL https://www.fda.gov/science-research/science-and-research-special-topics/artificial-intelligence-and-machine-learning-aiml-drug-development (accessed 10.8.24).U.S. Food & Drug Administration, 2023a. Center for Drug Evaluation and Research: Artificial Intelligence in Drug Manufacturing. U.S. Food & Drug Administration. https://doi.org/10.1016/j.ijpharm.2021.120554

U.S. Food & Drug Administration, 2023b. Emerging Technology Program (ETP) [WWW Document]. U.S. Food & Drug Administration. URL https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/emerging-technology-program-etp (accessed 10.8.24).

U.S. Food & Drug Administration, 2023c. FDA/PQRI Workshop on the Regulatory Framework for the Utilization of Artificial Intelligence in Pharmaceutical Manufacturing [WWW Document]. U.S. Food & Drug Administration. URL https://www.fda.gov/drugs/fdapqri-workshop-regulatory-framework-utilization-artificial-intelligence-pharmaceutical (accessed 10.8.24).U.S. Food & Drug Administration, 2021. Data Modernization Action Plan.

U.S. Food & Drug Administration, 2019. FDA's Technology Modernization Action Plan (TMAP).

U.S. Food & Drug Administration, Health Canada, Medicines & Healthcare products Regulatory Agency, 2024. Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles.U.S. Food & Drug Administration, Health Canada, Medicines & Healthcare products Regulatory Agency, 2023. Predetermined Change Control Plans for Machine Learning-Enabled Medical Devices: Guiding Principles.

U.S. Food & Drug Administration, Health Canada, Medicines & Healthcare products Regulatory Agency, 2021. Good Machine Learning Practice for Medical Device Development: Guiding Principles.

U.S. Food & Drugs Administration, 2024. FDA Digital Health and Artificial Intelligence Glossary – Educational Resource [WWW Document]. U.S. Food & Drugs Administration. URL https://www.fda.gov/science-research/artificial-intelligence-and-medical-products/fda-digital-health-and-artificial-intelligence-glossary-educational-resource (accessed 10.3.24).

Valette, E., Bril El-Haouzi, H., Demesure, G., 2023. Industry 5.0 and its technologies: A systematic literature review upon the human place into IoT- and CPS-based industrial systems. Comput Ind Eng 184. https://doi.org/10.1016/j.cie.2023.109426

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., Zhao, S., 2019. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. https://doi.org/10.1038/s41573-019-0024-5

Vassileiou, A.D., Robertson, M.N., Wareham, B.G., Soundaranathan, M., Ottoboni, S., Florence, A.J., Hartwig, T., Johnston, B.F., 2023. A unified ML framework for solubility prediction across organic solvents. Digital Discovery 2, 356–367. https://doi.org/10.1039/d2dd00024e

Veeva Systems Inc., 2025. Veeva Common Data Architecture for Life Sciences™ [WWW Document]. Veeva. URL https://www.veeva.com/cda/ (accessed 4.10.25).

Venkatasubramanian, V., Zhao, C., Joglekar, G., Jain, A., Hailemariam, L., Suresh, P., Akkisetty, P., Morris, K., Reklaitis, G. V., 2006. Ontological informatics infrastructure for pharmaceutical product development and manufacturing. Comput Chem Eng 30, 1482–1496. https://doi.org/10.1016/j.compchemeng.2006.05.036

Vijayakumar, A., Vairavasundaram, S., Koilraj, J.A.S., Rajappa, M., Kotecha, K., Kulkarni, A., 2024. Real-time visual intelligence for defect detection in pharmaceutical packaging. Sci Rep 14, 18811. https://doi.org/10.1038/s41598-024-69701-zViswanath, S., Guntz, S., Dieringer, J., Vaidyaraman, S., Wang, H., Gounaris, C., 2022. An Ontology to Describe Small Molecule Pharmaceutical Product Development and Methodology for Optimal Activity Scheduling. J Pharm Innov 17, 155–169. https://doi.org/10.1007/s12247-020-09505-6/Published

Voigt, S.P., Kalidindi, S.R., 2021. Materials graph ontology. Mater Lett 295. https://doi.org/10.1016/j.matlet.2021.129836

Walsche, R., 2021. The Road to Big Data Standardisation, in: Curry, E., Metzger, A., Zillner, S., Pazzaglia, J.-C., García Robles, A. (Eds.), The Elements of Big Data Value. Springer, Cham, Switzerland.

Wang, K., Han, L., Mustakis, J., Li, B., Magano, J., Damon, D.B., Dion, A., Maloney, M.T., Post, R., Li, R., 2020. Kinetic and Data-Driven Reaction Analysis for Pharmaceutical Process Development. Ind Eng Chem Res 59, 2409–2421. https://doi.org/10.1021/acs.iecr.9b03578

Wang, Z., Cao, J., Li, W., Wang, Y., Luo, G., Qiao, Y., Zhang, Y., Xu, B., 2021. Using a material database and data fusion method to accelerate the process model development of high shear wet granulation. Sci Rep 11. https://doi.org/10.1038/s41598-021-96097-x

Warzecha, M., Verma, L., Johnston, B.F., Palmer, J.C., Florence, A.J., Vekilov, P.G., 2020. Olanzapine crystal symmetry originates in preformed centrosymmetric solute dimers. Nat Chem 12, 914–920. https://doi.org/10.1038/s41557-020-0542-0

Waterman, K.C., MacDonald, B.C., 2010. Package selection for moisture protection for solid, oral drug products. J Pharm Sci. https://doi.org/10.1002/jps.22161

Werner, R., Takacs, R., Geier, D., Becker, T., Weißenberg, N., Haße, H., Sollacher, R., Thalhofer, M., Schumm, B., Steinke, I., 2021. The Challenge of Implementing Digital Twins in Operating Value Chains., in: Herwig, C., Pörtner, R., Möller, J. (Eds.), Digital Twins: Applications to the Design and Optimization of Bioprocesses. pp. 127–166.

Whatfix, 2024. What Is Digital Adoption? [WWW Document]. Whatfix. URL https://whatfix.com/digital-adoption/ (accessed 8.5.24).

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., t Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. Comment: The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3. https://doi.org/10.1038/sdata.2016.18

Wise, J., De Barron, A.G., Splendiani, A., Balali-Mood, B., Vasant, D., Little, E., Mellino, G., Harrow, I., Smith, I., Taubert, J., Van Bochove, K., Romacker, M., Walgemoed, P., Jimenez, R.C., Winnenburg, R., Plasterer, T., Gupta, V., Hedley, V., 2019. Implementation and relevance of FAIR data principles in biopharmaceutical R&D. Drug Discovery Today 24, 933–938. https://doi.org/10.1016/j.drudis.2019.01.008

Wu, J., Yang, X., Pan, Y., Zuo, T., Ning, Z., Li, C., Zhang, Z., 2023. Recent developments of automated flow chemistry in pharmaceutical compounds synthesis. J Flow Chem. https://doi.org/10.1007/s41981-023-00285-x

Wulf, C., Beller, M., Boenisch, T., Deutschmann, O., Hanf, S., Kockmann, N., Kraehnert, R., Oezaslan, M., Palkovits, S., Schimmler, S., Schunk, S.A., Wagemann, K., Linke, D., 2021. A Unified Research Data Infrastructure for Catalysis Research – Challenges and Concepts. ChemCatChem 13, 3223–3236. https://doi.org/10.1002/cctc.202001974

Xu, X., Lu, Y., Vogel-Heuser, B., Wang, L., 2021. Industry 4.0 and Industry 5.0—Inception, conception and perception. J Manuf Syst 61, 530–535. https://doi.org/10.1016/j.jmsy.2021.10.006

Yaacoub, J.P.A., Salman, O., Noura, H.N., Kaaniche, N., Chehab, A., Malli, M., 2020. Cyber-physical systems security: Limitations, issues and future trends. Microprocess Microsyst 77. https://doi.org/10.1016/j.micpro.2020.103201

Yang, C., Tomoshige, H., 2024. Self-Driving Labs: AI and Robotics Accelerating Materials Innovation [WWW Document]. Center for Strategic & International Studies. . URL https://www.csis.org/blogs/perspectives-innovation/self-driving-labs-ai-and-robotics-accelerating-materials-innovation (accessed 7.29.24).

Yang, X., Wang, Y., Byrne, R., Schneider, G., Yang, S., 2019. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. Chem Rev. https://doi.org/10.1021/acs.chemrev.8b00728

Yu, L.X., Amidon, G., Khan, M.A., Hoag, S.W., Polli, J., Raju, G.K., Woodcock, J., 2014. Understanding Pharmaceutical Quality by Design. AAPS J 16, 771–783. https://doi.org/10.1208/s12248-014-9598-3

Yu, L.X., Raw, A., Wu, L., Capacci-Daniel, C., Zhang, Y., Rosencrance, S., 2019. FDA's new pharmaceutical quality initiative: Knowledge-aided assessment & structured applications. Int J Pharm X 1. https://doi.org/10.1016/j.ijpx.2019.100010

Zaborenko, N., Shi, Z., Corredor, C.C., Smith-Goettler, B.M., Zhang, L., Hermans, A., Neu, C.M., Alam, M.A., Cohen, M.J., Lu, X., Xiong, L., Zacour, B.M., 2019. First-Principles and Empirical Approaches to Predicting In Vitro Dissolution for Pharmaceutical Formulation and Process Development and for Product Release Testing. AAPS Journal 21. https://doi.org/10.1208/s12248-019-0297-y

Zhang, L., Tan, J., Han, D., Zhu, H., 2017. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today. https://doi.org/10.1016/j.drudis.2017.08.010

Zhao, H., Yu, Y., Ni, N., Zhao, L., Lin, X., Wang, Y., Du, R., Shen, L., 2022. A new parameter for characterization of tablet friability based on a systematical study of five excipients. Int J Pharm 611. https://doi.org/10.1016/j.ijpharm.2021.121339

Zhou, N., Patel, D., Lin, S., O'donncha, F., 2024. Towards Automated Solution Recipe Generation for Industrial Asset Management with LLM A Preprint.

Figure 1: Illustration of QbDD scope adapted from (Su et al., 2019) showing the relationship between manufacturing technology development and strategies to define and control quality.

Figure 2: The transition from QbD to QbDD with reference to its effect on the knowledge space and the use of an existing data fabric (see Section 2.4.3) to inform experimentation and CPSs at each stage of development (as part of self-driving DataFactories (see Sections 2.4.1 and 2.4.3)) to enable a range of benefits. PP and MA refer to process parameters and material attributes, respectively.

Figure 3: A high level QbDD workflow is shown indicating the key stages where predictive tools will be required. For simplicity, this workflow does not show the feedback loops inherent in process development, but examples of feedback loops are discussed in this section.

Figure 4: A high-level diagram of elements interconnected and enabling the QbDD workflow.

Figure 5: A QbDD data fabric structures data and data handovers between the four CPRI platforms in this diagram. *Labels Skills, Measure, Model and Make, and letters (A)-(L) are described in Section 2.4.4.*

Highlights

- Quality by Digital Design (QbDD) enables a holistic end-to-end approach that ensures product quality and improves the sustainability of process and product development.
- QbDD can accelerate medicines development via the digital identification and exploration of more robust design spaces exploiting model-driven experimental approaches.
- The QbDD strategy uses workflows, ontologies, FAIR (findable, accessible, interoperable, and reusable) data and an underlying data fabric to reduce siloed data and enable digital predictive tools to span gaps in process and product design
- Collaborative academic and industry development efforts are needed to improve product, process, and performance predictive tools to fully exploit the benefits of QbDD.