MINI-REVIEW

Data integrity within the biopharmaceutical sector in the era of Industry 4.0

Haneen Alosert[1] | James Savery[2] | Jennifer Rheaume[3] | Matthew Cheeks[4] |
Richard Turner[5] | Christopher Spencer[4] | Suzanne S. Farid[1] | Stephen Goldrick[1]

[1]The Advanced Centre of Biochemical
Engineering, Department of Biochemical
Engineering, University College London,
London, UK

[2]Bioprocess Development Data Science &
Modelling, AstraZeneca, Cambridge, UK

[3]Equipment and Software Validation,
Development Quality Biologics, AstraZeneca,
Gaithersburg, Maryland, USA

[4]Cell Culture Fermentation Sciences,
Biopharmaceuticals Development R&D,
AstraZeneca, Cambridge, UK

[5]Purification Process Sciences,
Biopharmaceuticals Development R&D,
AstraZeneca, Cambridge, UK

Correspondence
Stephen Goldrick, The Advanced Centre of
Biochemical Engineering, Department of
Biochemical Engineering, University College
London, London, UK.
Email: s.goldrick@ucl.ac.uk

Abstract

Data Integrity (DI) in the highly regulated biopharmaceutical sector is of paramount importance to ensure decisions on meeting product specifications are accurate and hence assure patient safety and product quality. The challenge of ensuring DI within this sector is becoming more complex with the growing amount of data generated given increasing adoption of process analytical technology (PAT), advanced automation, high throughput microscale studies, and managing data models created by machine learning (ML) tools. This paper aims to identify DI risks and mitigation strategies in biopharmaceutical manufacturing facilities as the sector moves towards Industry 4.0. To achieve this, the paper examines common DI violations and links them to the ALCOA+ principles used across the FDA, EMA, and MHRA. The relevant DI guidelines from the ISPE's GAMP5 and ISA-95 standards are also discussed with a focus on the role of validated computerised and automated manufacturing systems to avoid DI risks and generate compliant data. The paper also highlights the importance of DI whilst using data analytics to ensure the developed models meet the required regulatory standards for process monitoring and control. This includes a discussion on possible mitigation strategies and methodologies to ensure data integrity is maintained for smart manufacturing operations such as the use of cloud platforms to facilitate the storage and transfer of manufacturing data, and migrate away from paper-based records.

KEYWORDS
biochemical engineering, bioprocess manufacturing, data analytics, data integrity, Industry 4.0, PAT

1 | INTRODUCTION

Product quality, safety, and efficacy are the biopharmaceutical industry's main concerns when manufacturing therapeutics. It can take over a decade to demonstrate these qualities through clinical trials and process development to finally achieve market authorization. Regulatory bodies need to review a significant amount of data to ensure good manufacturing practice (GMP) processes are robustly designed to consistently deliver high quality, efficacious, and safe products to patients. Regulators expect all product quality results to also meet the necessary data integrity (DI) standards.[1,2] Advanced technologies such as high throughput platforms and process analytical technologies (PAT) facilitate better process monitoring and control to ultimately assure improved product quality. These recent innovations have increased significantly the amount and complexity of the data generated during the manufacturing process. This rise in data complexity has encouraged the shift towards using more sophisticated methods to assist with decision-making, such as using statistical predictive models. The need to extract information from such complex datasets has further reinforced the criticality of DI in bioprocessing. DI is also integral to the success of Industry 4.0, which describes smart future factories that integrate autonomous real-time monitoring and control to enable improved responsiveness and flexibility.[3] Industry 4.0 relies on more investment into data management infrastructure such as data lakes or data warehouses that enable effective data storage and transfer to facilitate dissemination of information.[4] This review will address the criticality of DI and how to mitigate any potential DI risks within biomanufacturing processes that follow current good manufacturing practices (cGMP).

2 | DI STANDARDS

DI is defined as data that meets standards of completeness, accuracy and consistency. More specifically, in 2013 the FDA introduced the ALCOA acronym to indicate that data must be attributable, legible, contemporaneous, original and accurate, with the addition of being enduring and available which is commonly referred to as ALCOA+. Figure 1 summarizes and describes each ALCOA+ element to ensure data robustness and accessibility during data processing to help strengthen regulatory approvals for product applications.[5] Implementing ALCOA+ can help detect DI risks and avoid jeopardizing or delaying regulatory product approval which can lead to loss of processing time and material; hence ALCOA+ ultimately leads to cost savings.

3 | DI REGULATIONS AND VIOLATIONS

In recent years regulatory warning letters have risen, of which 43% featured DI issues that could jeopardize the regulatory status of companies.[6,7] In addition to the ALCOA+ principles, regulators such as the FDA, EMA, and MHRA have placed greater scrutiny on DI compliance and updated recent guidelines on the subject matter.[2,5,8–11]

Table 1 summarizes the key DI guidelines collated from the aforementioned regulators alongside common examples of DI violations and non-adherence to ALCOA+ principles in those instances. The guidelines span topics ranging from training staff on data processing through to checking internal audit trails and validating control strategies. Common DI violations relate typically to data manipulation and falsification. Examples include inattentive documentation leading to incomplete records due to insufficient training, for example, backdating batch records due to deficiencies in audit trails, through to intentional falsification of records to create acceptable results to fit target specification. For instance, in 2018 an Asian facility was warned for not complying to cGMP manufacturing regulations when HPLC data failed to meet specification and data was retested without providing clarity over why previous results failed to meet specifications leading to a DI violation.[1] This highlights the importance of sound scientific judgement needed to justify alterations made to restricted data and adhering to regulations on formal documents to record and modify procedures.

4 | MANUFACTURING SYSTEMS AND DI - RISKS AND MITIGATING MEASURES

Computerised systems consist of a controlling system (i.e., hardware, software, and firmware) and the network components that facilitate the control of a process such as creating, modifying and sharing digital information.[2] The international standard for control systems (ISA-95) was designed to define electronic information exchanged between the manufacturing control functions and other enterprise functions. Figure 2 adapts the ISA-95 levels by superimposing current examples of GxP bioprocessing data (i.e., GLP, GMP, and GDP where x = laboratory, manufacturing and documentation respectively) and instruments used for monitoring and control of a bioreactor. Level 0 describes a physical production process within a manufacturing environment.[12] Level 1 records processing data from instruments such as sensors or PAT tools. Typical data captured by these instruments in bioprocesses include online (e.g., pH or Raman), offline (e.g., titer), product quality (e.g., aggregates), and metadata (e.g., lot numbers). The data recorded feeds into real-time process monitoring and control seen in Level 2. This feeds into Level 3 on operational manufacturing, where for example manufacturing execution systems (MES) are used. Finally, the business planning level (Level 4) outlines the funds available and management of the systems and equipment designed to produce high-quality products to patients within the set timeframe.

Computerized systems pose some DI challenges, however, issues typically stem from inappropriate management of complex data records (e.g., PAT records) and failure to validate systems in use.[13] As more computerized systems are used by biopharmaceutical manufacturing facilities, there is a need to shift from a legacy paper-based approach to a fully electronic-based system to alleviate risks related to error prone tasks (e.g., manual pH records) and to streamline documentation.[13] To produce DI compliant data, software must be quality checked to verify and validate control strategies for their
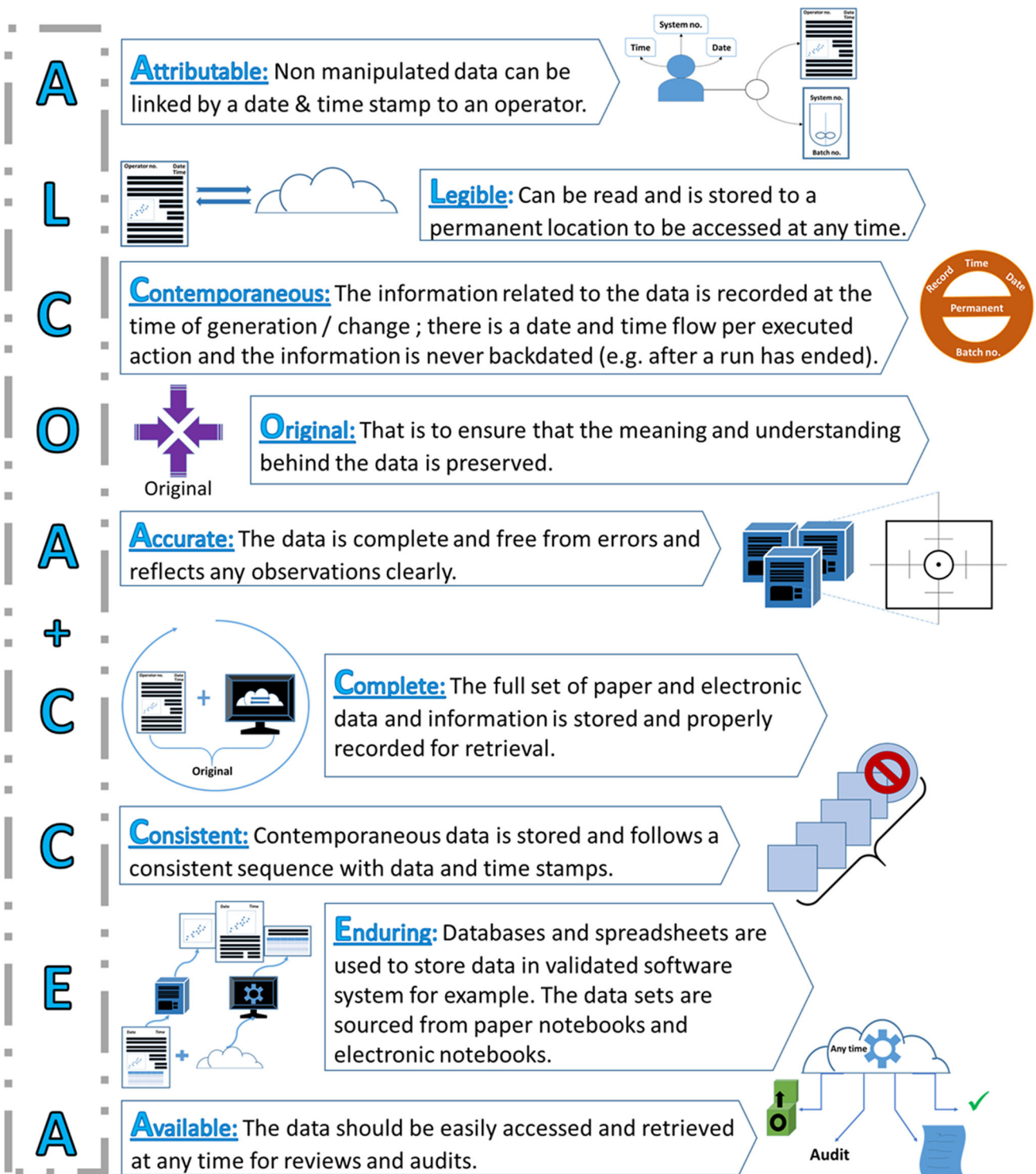
FIGURE 1 FDA's ALCOA+ principles broken down into its elements with their descriptions

intended use prior to their application in GxP environments.[4] Another method to reduce DI issues is through frequent internal audits and record reviews (e.g., system logs), which identify areas that fall short of DI compliance and enable mitigating measures.[11] Having a robust IT infrastructure can help ease the data review process and effectively manage data storage, transfer, and backup.[2,6] Hence, lacking a robust infrastructure poses risks to the continuity of a manufacturing process and ultimately the success of a product if a system breaks down.[14]

Good automated manufacturing practice (GAMP) was therefore established to provide a risk based approach for achieving compliant GxP computerised systems in industry, which includes meeting DI regulations.[2] GAMP5 (2008 update) in conjunction with ISA-95 can

TABLE 1 Common data integrity guidelines released by the EMA, FDA and MHRA shown against examples of data integrity violations and non-adherence to ALCOA+ standards

FDA, EMA, and MHRA- common data integrity guidelines	Common examples of data integrity violations	ALCOA+ related violations
Training all personnel on different data storage and processing formats preferably under Good Documentation Practice (GDP). This includes staff such as: • process operators • supervisors • quality assurance inspectors	• Inattentive documentation leading to potentially missing data during note taking or non-validated recording software • Data transferred from paper copies to electronic notebooks is not acceptable as it is defined as data manipulation	✗Complete ✗Accurate ✗Consistent
All data must be reviewed by QA departments, including: • computerised records stored in the cloud or shared drives • physical records	• Does not comply with the standards • Incomplete and inaccurately submitted documentation	✗Accurate ✗Complete ✗Enduring
All data forms must be recorded and stored safely as a backup for regulatory inspection, such as: • printed observations from analytical systems • raw non-processed data of electronic records • all meta-data recording in electronic notebooks (ELNs) and in lab books	• Data (raw / processed) unavailable for regulatory inspection • Manual observations recorded on loose paper • Archived copies are not an accurate representation of original records	✗Accurate ✗Available ✗Legible ✗Complete
Audit trails should be checked by QA departments for an accurate trail of data changes such as deletions and backdating. (This should be made available to regulatory inspectors for validation as well)	• Backdating results and observations • Back documenting activities • Creating acceptable results to fit with submitted specification validated by data deletion and data manipulation	✗Accurate ✗Complete ✗Contemporaneous ✗Enduring ✗Available
Changes to records must be restricted to personnel identifiable in a traceable manner such as through independent logins	• Using shared / common sign in in electronic systems that cannot be linked to a single operator / scientist with date and time stamps	✗Consistent ✗Attributable
Control strategies must be based on sound scientific basis and must be validated by QA personnel	• Measures in place to maintain and monitor data is non-validated	✗Accurate ✗Available

assist scientists and engineers who generate and interact with the data and systems. More documents such as GAMP's 2020 DI by design and the 2017 DI guides have also been released to provide further understanding and build better mitigating strategies related to DI risks.[15,16] Collectively, the guides clearly show the need for validated systems to generate compliant data necessary for process monitoring and control, in addition to needing GDP training for both manual and electronic transcripts.[8,9] GxP process training must involve a thorough explanation and link to regulations to create a company culture where DI issues are both disclosed and avoided, for example, using validated forms and templates.[9] Operators must also be provided with documentation such as standard operating procedures (SOP's) to ensure a consistent approach is carried out, data is managed appropriately and DI risks are reduced.[6] These guides allow scientists and quality staff to be more standardised through the creation of prototypes, specifications and action plans that expedite rapid and effective problem-solving and support.[6,8]

Processing facilities are comprised of a variety of computerised systems acquired by different vendors and working on different scales. Different analytical systems used would therefore store and generate data across different formats, such as CSV and txt files. The use of multiple and inconsistent data formats is a major limitation in the sec-

tor and more work is required to standardise these data formats.[4] To analyze these datasets, pre-processing and manipulation is required to produce readable results by an analyst. One solution is offered by the Allotrope Foundation that uses a standard data framework that facilitates the storage, use and integration of analytical data into a single file regardless of the data complexity to help avoid data loss and misinterpretation risks.[17] Such data frameworks also offer the integration and traceability of metadata (e.g., age of cell inoculum, cell type, and foaming issues) which can provide the necessary data context. Recording metadata is of paramount importance as it captures the essence and purpose of the experiments, simplifies analysis and helps leverage deeper understanding necessary for better process control and product quality assurance.

5 | INDUSTRY 4.0 - THE IMPACT OF DATA ANALYTICS AND SMART MANUFACTURING SOLUTIONS ON DI

Automation and digitalisation are pushing towards smart manufacturing solutions governed by Industry 4.0. These require enduring computerised systems to continue the manufacturing of high efficacy and
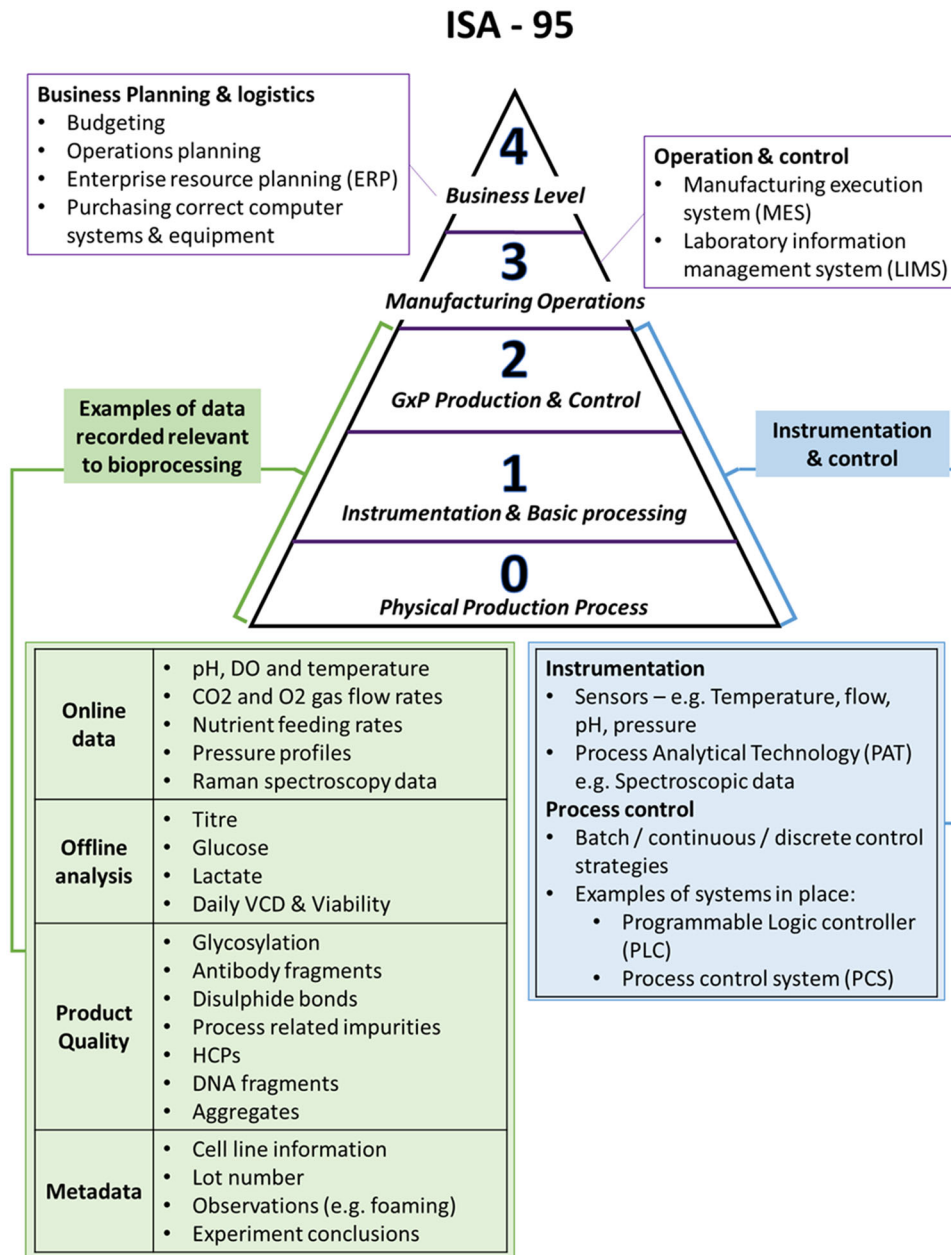
ISA - 95

Business Planning & logistics
- Budgeting
- Operations planning
- Enterprise resource planning (ERP)
- Purchasing correct computer systems & equipment

4 *Business Level*

Operation & control
- Manufacturing execution system (MES)
- Laboratory information management system (LIMS)

3 *Manufacturing Operations*

2 *GxP Production & Control*

Examples of data recorded relevant to bioprocessing

1 *Instrumentation & Basic processing*

Instrumentation & control

0 *Physical Production Process*

Online data	• pH, DO and temperature • CO2 and O2 gas flow rates • Nutrient feeding rates • Pressure profiles • Raman spectroscopy data	
Offline analysis	• Titre • Glucose • Lactate • Daily VCD & Viability	
Product Quality	• Glycosylation • Antibody fragments • Disulphide bonds • Process related impurities • HCPs • DNA fragments • Aggregates	
Metadata	• Cell line information • Lot number • Observations (e.g. foaming) • Experiment conclusions	

Instrumentation
- Sensors – e.g. Temperature, flow, pH, pressure
- Process Analytical Technology (PAT) e.g. Spectroscopic data

Process control
- Batch / continuous / discrete control strategies
- Examples of systems in place:
  - Programmable Logic controller (PLC)
  - Process control system (PCS)

FIGURE 2 ISA-95 five layer framework for computerized systems relevant to a GMP bioprocess manufacturing superimposed with a list of examples described under each respective layer

quality therapeutics.[4] The increased use of advanced PATs has further increased the complexity and volume of multidimensional data being generated, which require additional processing, model generation, and storage. Therefore, there is an increased need to use modelling tools such as machine learning (ML) including multivariate data analysis (MVDA), mechanistic models, and hybrid modelling techniques to leverage insights from this data. Likewise, all data analysis models and datasets processed must comply with DI standards.[18] Data generated is either recorded in its raw unedited form, such as raw CSV files, or in a processed form, such as data that has been restructured and fitted through a model. It is important to note that raw data must always be archived safely and made available for regula-

tory inspection during the validation period to ensure DI standards are met.[19] Therefore, the data and method of storage, transfer, and processing must be verified and documented to assure the data accuracy and integrity is preserved.[19] The use of PAT tools in conjunction with advanced ML techniques can be used to extract hidden information and acquire further understanding used for better monitoring and control.[20] The FDA encourages the implementation of these innovative tools and permits PAT data to be submitted in comparability protocols to validate processing implementation strategies and timelines.[21] Likewise, the regulatory concept of corrective and preventative action (CAPA) designed to identify, investigate and understand root causes of issues such as process deviations can be implemented to prevent
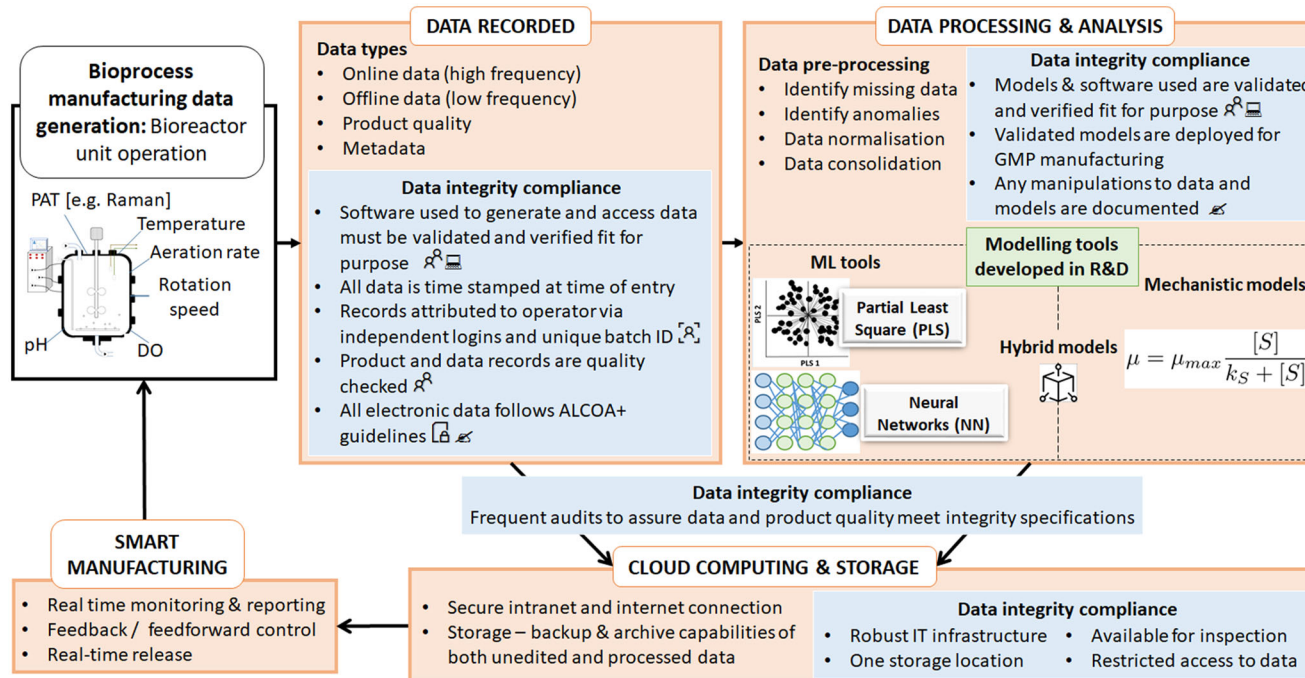
FIGURE 3 Necessary steps required to ensure data integrity is mantained in an Industry 4.0 bioprocess throughout a data lifecycle specific to a bioreactor. This encompasses the advanced technologies using smart manufacturing approaches to record, process and produce compliant results for bioprocess monitoring and control

reoccurrences.[21] GMP environments would benefit from integrating CAPA procedures to anticipate where processing may be compromised and lead to either costly delays or worse product rejections. Data analytics can be used to predict emerging problems through ML tools and advanced data analytics to estimate and recommend appropriate solutions, yet it is crucial to ensure that data processed meet DI standards.[20] Data analytics also has the potential to reduce deviations and failed batch runs to ultimately improve process control and shorten development timelines. Therefore, the use of data analytics at GMP level must be governed by models developed and validated in early RandD stages that have also met DI standards for commercial GMP manufacturing.[18,22] For this reason it is desired to ensure DI standards are met during early stages of development and maintained throughout scale up activities to minimise further DI risks in the GMP environment. A challenge with GMP data analysis is the type of data recorded, for example, time-series data such as those generated from online pH or DO sensors represent a form of high-frequency data acquisition; this normally requires complex dynamic models to extract useful correlations.[23] Ensuring that these complex datasets meet DI standards is therefore of particular importance. Data-driven smart manufacturing can help achieve effective data and resource management across different manufacturing sites. Smart manufacturing solutions such as cloud platforms can increase security and accessibility, to safely store and transfer large complex data volumes onto a single server to preserve its DI throughout processing.[24] For example, cloud based ELNs improve experimental logistic workflows and are within the same price range as managing manual logbooks.[25] This however requires a sophisticated built-in infrastructure to facilitate features

such as secure intranet and extranet, in addition to secure internet connection to communicate with trusted cloud providers.[24,26] The use of cloud platforms poses risks related to data transfer, data ownership and access, particularly in the context of global biopharmaceutical companies with multiple sites in various countries that may be governed by different data compliance regulations. One way to protect against data and system theft is having the appropriate cyber security measures setup against hacking and data theft.[6,14] This is vital as more remote access is made available for those working from home, thus requiring a two-point authentication to ensure a secure connection is made between the user and data source. Above all, the challenge with smart manufacturing is the availability of a fast and secure network connection to achieve the required short latency time to offer real-time monitoring capabilities.[20] This is crucial as there is a need to instantly access data stored in the cloud for effective real-time monitoring and control. Figure 3 summarises and illustrates the aforementioned solutions within Industry 4.0 that meet DI standards specific to a bioreactor. The bioreactor should be set up with verified and validated software, firmware and control strategies. In addition, SOPs are required, for example to calibrate probes and run maintenance that fulfils compliance requirements fit for process monitoring and control. All data captured such as standard sensor measurements (e.g., pH and DO), offline analytics and other advanced PAT sensor data (e.g., from Raman spectroscopy) are recorded and stored in their raw unedited form. A copy of the raw data is then pre-processed and handled using data analytics such as ML tools that have been developed during R&D and validated for monitoring and control at scale to then be deployed in a GMP manufacturing environment.[27] It is key to ensure all data and

model manipulations are recorded. To ensure DI compliance throughout and after the bioreactor run, records of frequent internal audits and logged errors must be stored appropriately. Once the data is stored and archived, online analysis can be performed by real-time monitoring and control strategies enabling smart manufacturing.

6 | FUTURE MEASURES TO MITIGATE DI RISKS

As the biopharmaceutical sector becomes more digitally mature and moves towards Industry 4.0, it is paramount to consider further measures to mitigate DI risks. Independent logins are suggested to act as identification signatures, even on shared systems to make the data attributable and traceable during regulatory auditing.[6] Future solutions also consider blockchain applications, which rely on multi-step verification of the data generated to ensure data traceability, transparency and security.[28,29] This method is a secure by design approach which allows manufacturing data to be stored within an incorruptible digital ledger with all corresponding transactions and relevant time stamps. Another solution is internal auditing that helps track procedures, action plans, and control measures implemented to determine if there is a need for requalification or flag DI violations early.[6] Internal audits will also facilitate preserving DI when using data analysis tools, likewise restricted access and irreversible recording methods will also help preserve DI particularly when needing to alter specifications and set process parameters.[5,11] Monitored connections can be also utilized to monitor use and track login which can also help trace anomalies and flag breaches to original datasets during inspection.[5,24] An electronic batch record or laboratory information management (LIMS) system is recommended to automatically save electronic entries.[5] Additionally, using numbered and controlled forms for manual transcripts recorded on portable tablets can also assist with quality checks, by avoiding for example loss of information due to bad handwriting.[11] Provided ALCOA+ principles are met, the industry is moving to more advanced autonomous recording systems such as using digital photos to the integration of lab voice systems. The aforementioned solutions are promising, specifically as it aids with documentation that allows for contemporaneous data recording.[10,27] These electronic records are however vulnerable to security breaches if not locked and controlled appropriately by using measures such as strict intranet sharing. The use of novel technologies in the near future built with the right infrastructure, storage server and relevant regulatory qualification can therefore alleviate some of the DI risks in GxP records.

7 | CONCLUSION

Complying with DI standards is a core part of the quality assurance procedure and helps biopharmaceutical companies ensure the continued product quality, efficacy and safety of their products. Standards such as the FDA's ALCOA+ guidelines have been released to help those in the biopharmaceutical sector to assess and mitigate DI risks and avoid costly regulatory product rejections. As the industry shifts towards Industry 4.0, the criticality of DI will be ever more paramount to enable the vision of smart factories that rely on digital integration and data analytics. This review has highlighted areas in R&D and GMP manufacturing that risk violating DI standards which include the use of more advanced tools such as PAT, data analytics, and cloud computing. The paper also suggests some mitigating measures to minimise and avoid the potential DI violations discussed.

CONFLICTS OF INTEREST
The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
Data sharing is not applicable as no new data was generated and/or the article describes entirely theoretical research and no new data.

ORCID
Haneen Alosert https://orcid.org/0000-0002-8070-2699
Suzanne S. Farid https://orcid.org/0000-0001-8155-0538
Stephen Goldrick https://orcid.org/0000-0002-8319-1679

REFERENCES
1. FDA. (2018, August 10). *Warning letter Kyowa Hakko Bio Co., Ltd*. U.S. FOOD & DRUG ADMINISTRATION. https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/kyowa-hakko-bio-co-ltd-543924-08102018
2. International Society for Pharmaceutical Engineering. (2008). *GAMP 5: A Risk-based Approach to Compliant Gxp Computerized Systems*. International Society for Pharmaceutical Engineering, North Bethesda, MD, USA. https://ispe.org/publications/guidance-documents/gamp-5
3. Branke, J., Farid, S. S., & Shah, N. (2016). Industry 4.0: A vision for personalized medicine supply chains? *Cell and Gene Therapy Insights*, 2(2), 263–270. https://doi.org/10.18609/cgti.2016.027
4. Steinwandter, V., Borchert, D., & Herwig, C. (2019). Data science tools and applications on the way to Pharma 4.0. *Drug Discovery Today*, 24(9), 1795–1805. https://doi.org/10.1016/j.drudis.2019.06.005
5. FDA. (2018). *Data Integrity and Compliance With Drug CGMP Questions and Answers Guidance for Industry* (Issue December). https://www.fda.gov/media/119267/download
6. Buytaert-Hoefen, K. (2019). A Harmonized approach to Data integrity. *BioProcess International*. https://bioprocessintl.com/manufacturing/information-technology/a-harmonized-approach-to-data-integrity/
7. Hodgson, D., Maini, F., Greenrose, W., Christiani, S., Chan, S., & Hargitai, B. (2017). Under the spotlight: Data integrity in life sciences. *Deloitte, London, UK*. https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Life-Sciences-Health-Care/gx-lshc-data-integrity.pdf
8. European commission. (2014). *The Rules Governing Medicinal Products in the European Union Volume 4 EU Guidelines for Good Manufactur-*

ing Practice for Medicinal Products for Human and Veterinary Use Part 1 Chapter 6: Quality Control. https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-4/2014-11_vol4_chapter_6.pdf

9. European Medicines Agency. (2016). Guidance on good manufacturing practice and good distribution practice. Questions and answers: Data integrity. https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-manufacturing-practice/guidance-good-manufacturing-practice-good-distribution-practice-questions-answers#data-integrity-(new-august-2016)-section

10. MHRA, G. (2018). Data integrity guidance and definitions. *Medicines and Healthcare products (Regulatory Agency, London,United Kingdom, 2018).* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/687246/MHRA_GxP_data_integrity_guide_March_edited_Final.pdf

11. Pérez, J. R. (2017). Maintaining data integrity. In *Quality Progress* (Vol. 50, Issue 3, pp. 14–15). http://bec-global.com/wp-content/uploads/2017/10/Article-Data-Integrity.pdf

12. Harjunkoski, I. (2014). Integration of scheduling and ISA-95. *Computer Aided Chemical Engineering, 33,* 427–432. https://doi.org/10.1016/B978-0-444-63456-6.50072-7

13. WHO. (2019). Guideline on data integrity. *WHO Drug Information, 33*(4), 773–793. http://www.who.int/medicines/areas/quality_safety/quality_assurance/guidelines/en/

14. European commission. (2011). Chapter 4: Documentation. In EudraLex The Rules Governing Medicinal Products in the European Union (Vol. 4, pp. 1–9). https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-4/chapter4_01-2011_en.pdf

15. ISPE. (2017). *GAMP Guide: Records & Data Integrity.* International Society for Pharmaceutical Engineering. https://ispe.org/publications/guidance-documents/gamp-records-pharmaceutical-data-integrity

16. International Society for Pharmaceutical Engineering. (2020). *GAMP RDI Good Practice Guide: Data Integrity by Design.* International Society for Pharmaceutical Engineering. https://ispe.org/publications/guidance-documents/gamp-rdi-good-practice-guide-data-integrity-design

17. Allotrope Foundation. (2017). *Creation of Allotrope Ontology and Usage as ELN Metadata* (Issue September). Allotrope Foundation. http://www.allotrope.orghttp//www.allotrope.org

18. Smiatek, J., Jung, A., & Bluhmki, E. (2020). Towards a digital bioprocess replica: Computational approaches in biopharmaceutical development and manufacturing. In *Trends in Biotechnology* (Vol. 38, Issue 10, pp. 1141–1153). Elsevier Current Trends. https://doi.org/10.1016/j.tibtech.2020.05.008

19. Römhild, A. (2016). Good manufacturing practice compliance in the manufacture of cell-based medicines. In *Guide to Cell Therapy GxP* (pp. 107–175). Elsevier Inc. https://doi.org/10.1016/b978-0-12-803115-5.00004-8

20. Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems, 48,* 157–169. https://doi.org/10.1016/j.jmsy.2018.01.006

21. Harris, J. R. (2008). Section 1 good manufacturing practices (GMP) and other FDA guidelines. In S. C. Gad (Ed.), *Pharmaceutical Manufacturing Handbook: Regulations and Quality* (pp. 1–43). John Wiley & Sons, Inc.

https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470259832.ch1?saml_referrer

22. von Stosch, M., Oliveira, R., Peres, J., & Feyo de Azevedo, S. (2014). Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers and Chemical Engineering, 60,* 86–101. https://doi.org/10.1016/j.compchemeng.2013.08.008

23. Goldrick, S., Holmes, W., Bond, N. J., Lewis, G., Kuiper, M., Turner, R., & Farid, S. S. (2017). Advanced multivariate data analysis to determine the root cause of trisulfide bond formation in a novel antibody–peptide fusion. *Biotechnology and Bioengineerng, 114*(10), 2222–2234. https://doi.org/10.1002/bit.26339

24. Lu, Y., & Xu, X. (2019). Cloud-based manufacturing equipment and big data analytics to enable on-demand manufacturing services. *Robotics and Computer-Integrated Manufacturing, 57,* 92–102. https://doi.org/10.1016/j.rcim.2018.11.006

25. Riley, M., Hattaway, E. Z., & Arthur Felse, P. (2017). Implementation and use of cloud-based electronic lab notebook in a bioprocess engineering teaching laboratory. *Journal of Biological Engineering, 11*(1), 1–9. https://doi.org/10.1186/s13036-017-0083-2

26. Paldy, E., & Densow, H. (2021). *Data integrity and digital workflow guidance in a contemporary pharma lab.* European Pharmaceutical Review. https://www.europeanpharmaceuticalreview.com/webinar/139335/data-integrity-and-digital-workflow-guidance-in-a-contemporary-pharma-lab/2/

27. Gargalo, C. L., Udugama, I., Pontius, K., Lopez, P. C., Nielsen, R. F., Hasanzadeh, A., Mansouri, S. S., Bayer, C., Junicke, H., & Gernaey, K. V. (2020). Towards smart biomanufacturing: A perspective on recent developments in industrial measurement and monitoring technologies for bio-based production processes. *Journal of Industrial Microbiology and Biotechnology, 47*(11), 947–964. https://doi.org/10.1007/s10295-020-02308-1

28. Leal, F., Chis, A. E., Caton, S., González–Vélez, H., García–Gómez, J. M., Durá, M., Sánchez–García, A., Sáez, C., Karageorgos, A., Gerogiannis, V. C., Xenakis, A., Lallas, E., Ntounas, T., Vasileiou, E., Mountzouris, G., Otti, B., Pucci, P., Papini, R., Cerrai, D., & Mier, M. (2021). Smart pharmaceutical manufacturing: Ensuring end-to-end traceability and data integrity in medicine production. *Big Data Research, 24,* 100172. https://doi.org/10.1016/j.bdr.2020.100172

29. Hemphill, T. A. (2020). Biologics regulation, second-to-market competition, and the use of blockchain technology: An opportunity for the FDA to support responsible biotechnology innovation. *Journal of Responsible Innovation, 7*(3), 689–696. https://doi.org/10.1080/23299460.2020.1807669