FLOW IMAGING MICROSCOPY AND MACHINE LEARNING METHODS AND APPLICATIONS FOR PARTICLE MORPHOLOGY ANALYSIS

By

AUSTIN LEWIS DANIELS

B.S., Auburn University, 2015

A thesis submitted to the
faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Chemical and Biological Engineering
2021

Committee Members:

Dr. Theodore W. Randolph

Dr. Christopher P. Calderon

Dr. Daniel K. Schwartz

Dr. Michael R. Shirts

Dr. John F. Carpenter

Abstract

Daniels, Austin Lewis (Ph.D., Chemical and Biological Engineering)

Flow imaging microscopy and machine learning methods and applications for particle morphology analysis

Thesis directed by Gilespie Professor Theodore W. Randolph and Associate Professor Adjunct Christopher P. Calderon

Flow imaging microscopy (FIM) is an increasingly popular technique for collecting light microscopy images of particles larger than 1 μm in samples such as therapeutic protein formulations. While FIM is commonly used to monitor the number of particles in a sample, the particle images returned by FIM contain particle morphology information that may be used to distinguish between different types of particles. For example, this morphology information may be used to differentiate between protein aggregates formed by different mechanisms or identify cells of different species. However, this analysis has previously been prohibited by the difficulty of extracting and analyzing relevant particle morphology information from FIM images.

This thesis describes the development of algorithms that use statistical and machine learning methods to analyze particle morphology information in FIM images and applications of these algorithms in protein formulation development. Several algorithms were developed that use convolutional neural networks (ConvNets) and other approaches to analyze the morphology of both individual particles as well as the particle population in a sample. These algorithms were then used to identify the impact of conditions such as accelerated and real-time stability stresses on protein aggregate morphology. This thesis will also describe generalizations of this approach for analyzing light and fluorescence microscopy images collected from imaging flow cytometry (IFC) and for distinguishing between blood cells and cells of different bacterial species as proof-of-concept for a FIM-based bloodstream infection (BSI) diagnostic test. These results demonstrate the effectiveness and usefulness of FIM-based particle morphology analysis techniques for analyzing particle morphology in therapeutic protein formulations, blood samples, and other particle-containing samples.

Acknowledgements

Many of the studies described in this thesis were performed in collaboration with other researchers both within and outside of the University of Colorado. These collaborators generated samples and/or data that were analyzed using the methods described in this manuscript as well as assisted with the experimental design. The flow imaging microscopy data presented in Chapter 3 was collected by the FDA (see Kotarek *et al.* 2016) and was obtained via a freedom of information act request. Sample generation and imaging described in Chapter 4 was performed with the assistance of Braden Carroll, Chanel Hill, and Kiersten Johnson. The study design, sample generation and imaging described in Chapters 5 and 6 was performed by Alyssa Witeof and Dr. Sanli Movafaghi, respectively. Samples analyzed in Chapter 7 were prepared with the assistance of Channel Hill. Imaging flow cytometry data presented in Chapter 7 was collected by Dr. Christine Probst at Luminex Co. / KBI Biopharma. Kristen Eller, Natalie Meinerz, Kathryne Walker, Hannah Middlestead, and Alyssa Witeof assisted with sample preparation in the results described in Chapter 8. Dr. Saba Ghazvini and Dr. Nicole Wallace assisted with the sample preparation and analysis described in Appendix A.

Funding for the work described in this thesis was supported by AstraZeneca, Yokogawa Fluid Imaging Technologies and Ursa Analytics. This research was also funded by NIH grant RO1 EB006006, NIH award R41GM130513, and Department of Education GAANN Award P200A150211

Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Protein Formulations and Aggregation

Therapeutic protein formulations are among the fastest growing classes of pharmaceuticals. When properly designed, these protein formulations can mimic the functions of native human proteins and can be used to replace (e.g. insulin replacement for diabetes patients) or augment (e.g. monoclonal antibodies targeted against pathogens like the COVID-19 virus) the function of native human proteins. However, the limited stability of proteins in aqueous conditions[1] poses challenges in designing and manufacturing stable, effective protein formulations. This instability often results in the formation of aggregates both under quiescent conditions and in response to stresses these formulations experience during manufacturing and development. These stresses include agitation[2], air-water and container-water interfaces[3–5], extreme pH[1] and freeze-thaw cycles[6–8] as well as combinations of these stresses[8]. Thus, many routine processes that these formulations undergo such as filtration[9], fill-finish[10–12], and shipping[13] can trigger aggregation. Other formulation components such as the container-closure system can also contribute particles to the formulation (e.g. glass flakes from glass containers[14], silicone oil microdroplets from prefilled syringes[15]). Virtually all therapeutic protein formulations contain particles due to these varied and often ubiquitous aggregate and particle sources.[16]

Strategies for monitoring protein aggregates and other particles in protein formulations are of interest both to companies and regulatory agencies like the FDA. This interest is due in large part to regulations such as USP <788> and <789> that require companies to monitor the particle content of these formulations to the best of their ability while also setting firm limits on the number of visible particles and those larger than 10 and 25 μm that can be present in the final formulation. Particle monitoring not only is essential to ensure formulations meet these regulations but may be useful in designing formulations and manufacturing processes that effectively control the particle content they generate. Particle monitoring strategies may also be essential to control the risk of adverse responses that these particles may influence. Both *in vitro* studies[17–20] and clinical experiences[21,22] have indicated that changes in particle content between samples are correlated with changes in adverse responses to the formulation including patient fatalities[22]. It is currently unknown if the relationship between particle content and adverse drug responses is causal and, if so, what features of these particles results in elevated immunogenicity. Until such information is available strategies for monitoring and controlling the particle content in a formulation would minimize the potential risk of adverse responses these particles pose.

1.2 Flow Imaging Microscopy (FIM) and Particle Imaging

Flow imaging microscopy (FIM) is an increasingly popular technique for imaging subvisible particles or particles larger than 1 μm in protein formulations[23,24]. FIM instruments such as FlowCam (Yokogawa Fluid Imaging Technologies, Inc, Scarborough, ME.) or MFI (ProteinSimple, San Jose, CA)[24] flow a sample through a small microfluidics channel and use light microscopy to image each particle in the sample with higher throughput than typical light microscopy. This analysis typically returns $>10^3$ images per milliliter of sample analyzed each

typically containing only a single particle. These images will be referred to as “FIM images” or, since there is usually only one particle per FIM image, as “particles” in the context of FIM data analysis. The FIM images from a sample can be counted to determine the number and concentration of particles. These images are also rich in morphology information about the particles which can be analyzed to determine the types of particles in a sample[25–28].

FIM has been increasingly used to monitor the particle content in therapeutic protein formulations. One common use for FIM is to monitor particle generation in accelerated stability studies, studies in which protein formulations are exposed to conditions designed to accelerate aggregation in a manner that simulates common aggregation-inducing conditions[17,29]. FIM can be used in these analyses to monitor the level and potentially the types of particles formed in different protein formulations (e.g. different containers) or by different stresses[30,31]. FIM has also shown promise for detecting changes in particle content in formulations due to either intentional (e.g. change in primary packaging type) or unintentional (e.g. process upset) deviations in manufacturing conditions[31].

The success of FIM in protein formulation applications has also generated interest in developing new and/or adapting existing particle imaging modalities to analyze particles in protein formulations. FlowCam Nano instruments operates in a similar fashion to standard FIM instruments but uses immersion oil-based optics to obtain higher particle magnification, enabling the instrument to image particles between 300 nm – 1 µm typically too small to detect via FIM. Imaging flow cytometry (IFC) obtains fluorescence microscopy images along with the standard light microscopy images collected by FIM[32]. Holographic video microscopy collects holographs of each particle which can be used to obtain 3D reconstructions of particle morphology as well as accurate measurements of particle size and refractive index[33,34]. Background membrane imaging uses membrane microscopy to mount images in a solvent-free environment before

image collection[35], resulting in more effective imaging of translucent particles that can be nearly invisible in FIM images collected in solvent[36]. While each of these techniques improves on the particle information that can be collected from FIM, standard FIM image analysis is more commonly used in practice and will be the focus of the subsequent discussion.

1.3 FIM and Particle Morphology

While FIM is increasingly used to monitor the particle content in formulations, most researchers in pharmaceuticals use FIM to measure particle concentrations. For example, researchers performing accelerated stability studies to compare formulations often use FIM to determine which formulation generated the fewest particles under accelerated stress conditions in hopes that the formulation will also mitigate particle formation under more realistic storage and stress conditions[30,31,37,38]. However, the particle images returned by FIM may also contain information about the particle types and morphologies in the sample. As protein aggregates and other particles can be made by dramatically different mechanisms[39–41], it is likely that different mechanisms can produce particles detectable via FIM that exhibit morphologies characteristic of that mechanism. If so, the particle morphologies present in FIM images can act as “fingerprints” of these different mechanisms of particle formation—potentially relevant information about the formulation and any damaging conditions it was exposed to. For example, this information could be used to identify the sources of an unexpected particle type in a formulation such as those that might be generated by a process upset. This information could also be used to determine if a change in formulation or process conditions such as formulating a protein in a different primary container induces a change in aggregation mechanisms formed during stresses encountered during shipping and storage—a potentially relevant difference in product stability.

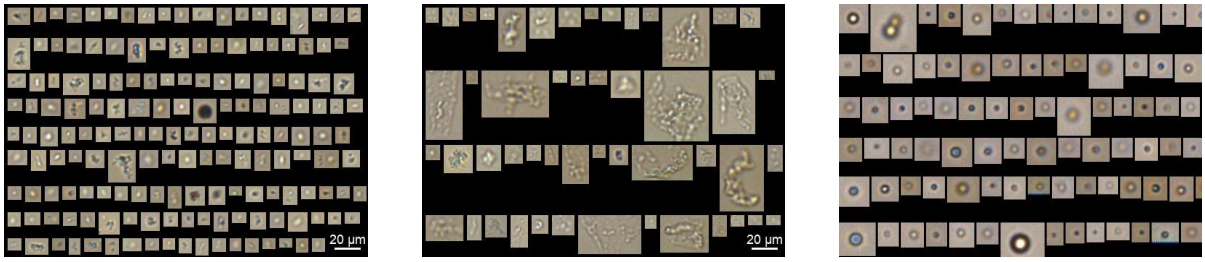

Figure 1: Sample Flow Imaging Microscopy (FIM) images
FIM images of protein aggregates made by freeze-thaw stress (left) and shaking stress (middle) as well as silicone oil microdroplets (right, not to scale)

While the discussion so far has centered on analyzing protein formulations via FIM, other samples contain particles that can be analyzed via FIM. For example, FlowCam instruments were initially used by water quality researchers to analyze algae in water[25,42] and have seen use for analyzing cell-based medicinal products[27]. It is easy to envision applications for these morphology analysis tools when analyzing other types of particles. For samples containing cells, these morphology tools could be used to identify different cell types in a sample such as identifying a contaminating microorganism in pond water or detecting rare cell types like bacteria or cancer cells in a blood sample to diagnose diseases.

1.4 Prior FIM Particle Morphology approaches

Unfortunately, analyzing particle morphology using FIM instruments has historically been difficult for researchers due in part to the difficulty of analyzing FIM image data. Typical FIM images can be thought of as 20 x 20-pixel images with either one or three channels per pixel for grayscale or color images respectively. As these images are effectively 1,200-dimension vectors, it is difficult to analyze all the information encoded in the raw image directly especially since individual pixels do not typically contain significant particle morphology information. To

make this analysis feasible, most image analyses will first develop a set of image features for raw FIM images that capture the most relevant information from the raw FIM image in a much lower dimensional representation. These lower-dimensional representations are more likely to be discriminative of different particles types than the raw pixel data and can then be more easily analyzed using standard statistics or machine learning tools. The effectiveness of any particle morphology analysis scheme thus depends both on the quality of the image features and the algorithm used to analyze particles in terms of these features.

Previous work used combinations of standard particle properties like particle size and aspect ratio as image features when trying to distinguish between different particle types. Most prior work in this area has focused on differentiating FIM images of protein aggregates from those of simple, spherical particles like silicone oil microdroplets[26,43,44] or air bubbles[23]. While these studies have used increasingly complex machine learning tools to perform this classification ranging from simple property thresholds[43] to random forest classifiers[26], each of these methods used the simple particle properties returned by FIM instruments as image features and typically focused on those that can be used to detect circular particle morphologies. A similar property-based approach was also previously used to distinguish between viable and unviable mammalian cells[27].

While these particle properties are useful for describing simple particle morphologies, these properties bias subsequent analysis towards changes in the simple particle morphology features these properties describe. As it is unlikely that these properties fully capture the relevant particle morphology information encoded in the image, these approaches may miss changes in complex particle morphology features that are not captured in these simple particle properties. It is desirable to develop image features that more effectively capture complex particle morphologies than is possible using simple particle properties. It would also be ideal for these image features

to be developed in an automated fashion. While manual feature engineering is possible, the heterogeneous nature of FIM image sets (see Figure 1) makes it difficult to manually identify and quantify the relevant morphology features present in raw FIM images.

Another key issue with these previous techniques for analyzing protein formulations is their focus on analyzing the morphology of individual particles rather than that of the particle population in a sample. Protein formulations encountered in practice may exhibit numerous particle morphologies due to the many aggregation-inducing stresses these formulations can be exposed to. While identifying individual particle morphologies can be useful in some applications, in pharmaceutical applications we are often more interested in analyzing the "particle morphology distribution" or the types and relative amounts of different particle morphologies in a sample. Strategies for analyzing particle populations could be used to compare aggregate populations between different samples of a formulation more effectively than techniques for analyzing individual particles.

1.5 Thesis Overview

This thesis describes my work in developing and applying machine learning tools to extract and analyze particle morphology information found in FIM measurements. These techniques use machine learning and statistical tools to either identify different types of particles in a sample or to check if a sample contains a particle population consistent with that in another sample. While much of the work involved developing and implementing these approaches, this work also focused on demonstrating how these techniques can be useful to researchers working with FIM or similar imaging platforms. Most of these applications were focused on protein formulations and how these techniques may be used to identify changes in protein aggregate morphology

between formulations or to identify potential sources of those aggregates. However, we also show how similar techniques may be useful for differentiating between different species of cells and how this approach may be useful for applications such as diagnosing sepsis. We also demonstrate how the approaches developed using FIM can also be generalized to other imaging modalities such as imaging flow cytometry that extract additional morphology information through orthogonal measurements that can be included in the analysis.

The remaining sections of the thesis are organized as follows. Chapters 2 discusses the evolution of the algorithm from a more sophisticated technique to analyzing common particle properties to convolutional neural network (ConvNet)-based strategies to learn more sophisticated and useful image features. Chapters 3 and 4 describes two of these algorithms in more detail as well as some of the initial applications that were pursued when developing these algorithms. Chapters 5 and 6 focus on applications of these algorithms to analyze aggregation caused by shipping stress (Chapter 5) and severe mechanical stresses and cavitation (Chapter 6). Chapters 7 and 8 discuss generalizations of these algorithms to imaging flow cytometry and sepsis diagnosis, respectively. Lastly, chapter 9 summarizes the findings described elsewhere in the thesis.

1.6 References

1. Chi EY, Krishnan S, Randolph TW, Carpenter JF. Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharm Res*. 2003;20(9):1325-1336. doi:10.1023/A:1025771421906

2. Kiese S, PappenBerger A, Friess W, Mahler H-C. Shaken, Not Stirred: Mechanical Stress Testing of an IgG1 Antibody. *J Pharm Sci*. 2008;97(10):4347-4366. doi:10.1002/jps

3. Vörös J. The density and refractive index of adsorbing protein layers. *Biophys J*.

2004;87(1):553-561. doi:10.1529/biophysj.103.030072

4. Bee JS, Randolph TW, Carpenter JF, Bishop SM, Dimitrova MN. Effects of surfaces and leachables on the stability of biopharmaceuticals. *J Pharm Sci*. 2011;100(10):4158-4170. doi:10.1002/jps.22597

5. Bee JS, Schwartz DK, Trabelsi S, et al. Production of particles of therapeutic proteins at the air-water interface during compression/dilation cycles. *Soft Matter*. 2012;8(40):10329-10335. doi:10.1039/c2sm26184g

6. Bhatnagar BS, Bogner RH, Pikal MJ. Protein stability during freezing: Separation of stresses and mechanisms of protein stabilization. *Pharm Dev Technol*. 2007;12(5):505-523. doi:10.1080/10837450701481157

7. Barnard JG, Singh S, Randolph TW, Carpenter JF. Subvisible particle counting provides a sensitive method of detecting and quantifying aggregation of monoclonal antibody caused by freeze-thawing: insights into the roles of particles in the protein aggregation pathway. *J Pharm Sci 2011*. 2011;100(2):492-503.

8. Pardeshi NN, Zhou C, Randolph TW, Carpenter JF. Protein Nanoparticles Promote Microparticle Formation in Intravenous Immunoglobulin Solutions During Freeze-Thawing and Agitation Stresses. *J Pharm Sci*. 2018;107(7):1852-1857. doi:10.1016/j.xphs.2018.03.016

9. Barnard JG, Kahn D, Cetlin D, Randolph TW, Carpenter JF. Investigations into the fouling mechanism of parvovirus filters during filtration of freeze-thawed mAb drug substance solutions. *J Pharm Sci*. 2014. doi:10.1002/jps.23881

10. Nayak A, Colandene J, Bradford V, Perkins M. Characterization of subvisible particle formation during the filling pump operation of a monoclonal antibody solution. *J Pharm Sci*. 2011;100:4198-4204.

11. Her C, Carpenter JF. Effects of Tubing Type, Formulation, and Postpumping Agitation on Nanoparticle and Microparticle Formation in Intravenous Immunoglobulin Solutions Processed With a Peristaltic Filling Pump. *J Pharm Sci*. 2020;109(1):739-749. doi:10.1016/j.xphs.2019.05.013

12. Wu H, Randolph TW. Aggregation and Particle Formation During Pumping of an Antibody Formulation Are Controlled by Electrostatic Interactions Between Pump Surfaces and Protein Molecules. *J Pharm Sci*. 2020;109(4):1473-1482. doi:10.1016/j.xphs.2020.01.023

13. Siska C, Harber P, Kerwin BA. Shocking Data on Parcel Shipments of Protein Solutions. *J Pharm Sci*. 2020;109:690-695.

14. Ennis RD, Pritchard R, Nakamura C, et al. Glass vials for small volume parenterals: Influence of drug and manufacturing processes on glass delamination. *Pharm Dev Technol*. 2001;6(3):393-405. doi:10.1081/PDT-100002248

15. Gerhardt A, McGraw NR, Schwartz DK, Bee JS, Carpenter JF, Randolph TW. Protein aggregation and particle formation in prefilled glass syringes. *J Pharm Sci*. 2014;103(6):1601-1612. doi:10.1002/jps.23973

16. Singh SK, Afonina N, Awwad M, et al. An industry perspective on the monitoring of subvisible particles as a quality attribute for protein therapeutics. *J Pharm Sci*. 2010;99(8):3302-3321.

17. Joubert MK, Luo Q, Nashed-Samuel Y, Wypych J, Narhi LO. Classification and characterization of therapeutic antibody aggregates. *J Biol Chem*. 2011;286(28):25118-25133. doi:10.1074/jbc.M110.160457

18. Joubert MK, Hokom M, Eakin C, et al. Highly aggregated antibody therapeutics can enhance the in vitro innate and late-stage T-cell immune responses. *J Biol Chem*. 2012;287(30):25266-25279. doi:10.1074/jbc.M111.330902

19. Chisholm CF, Baker AE, Soucie KR, Torres RM, Carpenter JF, Randolph TW. Silicone Oil Microdroplets Can Induce Antibody Responses Against Recombinant Murine Growth Hormone in Mice. *J Pharm Sci*. 2016;105(5):1623-1632. doi:10.1016/j.xphs.2016.02.019

20. Chisholm CF, Soucie KR, Song JS, et al. Immunogenicity of Structurally Perturbed Hen Egg Lysozyme Adsorbed to Silicone Oil Microdroplets in Wild-Type and Transgenic Mouse Models. *J Pharm Sci*. 2017;106(6):1519-1527.

21. Rosenberg AS. Effects of protein aggregates: An immunologic perspective. *AAPS J*. 2006;8(3):E501-E507. doi:10.1208/aapsj080359

22. Kotarek J, Stuart C, De Paoli SH, et al. Subvisible Particle Content, Formulation, and Dose of an Erythropoietin Peptide Mimetic Product Are Associated with Severe Adverse Postmarketing Events. *J Pharm Sci*. 2016;105(3):1023-1027. doi:10.1016/S0022-3549(15)00180-X

23. Sharma DK, King D, Oma P, Merchant C. Micro-Flow Imaging: Flow Microscopy Applied to Sub-visible Particulate Analysis in Protein Formulations. *AAPS J*. 2010;12(3):455-464. doi:10.1208/s12248-010-9205-1

24. Zölls S, Weinbuch D, Wiggenhorn M, et al. Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments. *AAPS J*. 2013;15(4):1200-1211. doi:10.1208/s12248-013-9522-2

25. Buskey EJ, Hyatt CJ. Use of the FlowCAM for semi-automated recognition and enumeration of red tide cells (Karenia brevis) in natural plankton samples. *Harmful Algae*. 2006;5(6):685-692. doi:10.1016/j.hal.2006.02.003

26. Saggu M, Patel AR, Koulis T. A Random Forest Approach for Counting Silicone Oil Droplets and Protein Particles in Antibody Formulations Using Flow Microscopy. *Pharm Res*. 2017;34(2):479-491. doi:10.1007/s11095-016-2079-x

27. Sediq AS, Klem R, Nejadnik MR, Meij P, Jiskoot W. Label-Free, Flow-Imaging Methods for Determination of Cell Concentration and Viability. *Pharm Res*. 2018;35(8). doi:10.1007/s11095-018-2422-5

28. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

29. Gambe-Gilbuena A, Shibano Y, Krayukhina E, Torisu T, Uchiyama S. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images. *J Pharm Sci*. 2020;109(1):614-623. doi:10.1016/j.xphs.2019.10.034

30. Narhi LO, Corvari V, Ripple DC, et al. Subvisible (2-100 μm) particle analysis during biotherapeutic drug product development: Part 1, considerations and strategy. *J Pharm Sci*. 2015;104(6):1899-1908. doi:10.1002/jps.24437

31. Corvari V, Narhi LO, Spitznagel TM, et al. Subvisible (2-100 μm) particle analysis during biotherapeutic drug product development: Part 2, experience with the application of subvisible particle analysis. *Biologicals*. 2015;43(6):457-473. doi:10.1016/j.biologicals.2015.07.011

32. Probst C. Characterization of Protein Aggregates, Silicone Oil Droplets, and Protein-Silicone Interactions Using Imaging Flow Cytometry. *J Pharm Sci*. 2019:1-11. doi:10.1016/j.xphs.2019.05.018

33. Wang C, Zhong X, Ruffner DB, et al. Holographic Characterization of Protein Aggregates. *J Pharm Sci*. 2016;105(3):1074-1085. doi:10.1016/j.xphs.2015.12.018

34. Kasimbeg PNO, Cheong FC, Ruffner DB, Blusewicz JM, Philips LA. Holographic Characterization of Protein Aggregates in the Presence of Silicone Oil and Surfactants. *J Pharm Sci*. 2018;108(1):155-161. doi:10.1016/j.xphs.2018.10.002

35. Helbig C, Ammann G, Menzen T, Friess W, Wuchner K, Hawe A. Backgrounded Membrane Imaging (BMI) for High-Throughput Characterization of Subvisible Particles During Biopharmaceutical Drug Product Development. *J Pharm Sci*. 2019:1-13. doi:10.1016/j.xphs.2019.03.024

36. Zölls S, Gregoritza M, Tantipolphan R, et al. How Subvisible Particles Become Invisible — Relevance of the Refractive Index for Protein Particle Analysis. *Pharm Biotechnol*. 2013;102(5):1434-1446. doi:10.1002/jps

37. Teska BM, Brake JM, Tronto GS, Carpenter JF. Aggregation and Particle Formation of Therapeutic Proteins in Contact With a Novel Fluoropolymer Surface Versus Siliconized Surfaces: Effects of Agitation in Vials and in Prefilled Syringes. *J Pharm Sci*. 2016;105(7):2053-2065. doi:10.1016/j.xphs.2016.04.015

38. Wu H, Chisholm CF, Puryear M, et al. Container Surfaces Control Initiation of Cavitation and Resulting Particle Formation in Protein Formulations After Application of Mechanical Shock. *J Pharm Sci*. 2020;109(3):1270-1280. doi:10.1016/j.xphs.2019.11.015

39. Roberts CJ. Non-native protein aggregation kinetics. *Biotechnol Bioeng*. 2007;98(5):927-938. doi:10.1002/bit.21627

40. Roberts CJ. Therapeutic protein aggregation: Mechanisms, design, and control. *Trends Biotechnol*. 2014;32(7):372-380. doi:10.1016/j.tibtech.2014.05.005

41. Wang L, Li Y, Huang J, Lazebnik S. Learning Two-Branch Neural Networks for Image-Text Matching Tasks. *IEEE Trans Pattern Anal Mach Intell*. 2018:1-14. doi:10.1109/TPAMI.2018.2797921

42. Sieracki CK, Sieracki ME, Yentsch CS. An imaging-in-flow system for automated analysis of marine microplankton. *Mar Ecol Prog Ser*. 1998;168:285-296. doi:10.3354/meps168285

43. Strehl R, Rombach-Riegraf V, Diez M, et al. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: A novel multiparametric image filter for sub-visible particles in microflow imaging analysis. *Pharm Res*. 2012;29(2):594-602. doi:10.1007/s11095-011-0590-7

44. Vandesteeg N, Kilbert C. Differentiation of Subvisible Silicone Oil Droplets from Irregular Standard Dust Particles. *J Pharm Sci*. 2013;10:1696–1700. doi:10.1002/jps

Chapter 2: Algorithm Development

Sections of this chapter based on Maddux et al. 2017[1], Calderon et al. 2018[2], and Daniels et al. 2020[3]. Full citations can be found in the references section of this chapter

2.1 Overview

Several algorithms were developed for analyzing particle morphology information embedded in FIM datasets. These algorithms were designed to address one of two core problems: particle classification and particle population analysis.

In particle classification problems, the goal was to distinguish between several known particle types similar to previous silicone oil detection approaches[4–7]. These algorithms were trained on labelled FIM images of different particle types to learn the particle morphology features that distinguished these particle types. The trained algorithm was then applied to new, unlabeled FIM images to predict which of those particle types was present in the image. One possible application for this approach is in root cause analysis; this approach could be used to recognize protein aggregates generated by known stresses and process upsets to help identify the root cause of a process upset.

In particle population analysis problems, the goal was to instead check if the particle morphology distribution or the types and relative amounts of different particle morphologies in a sample are consistent with the distribution in a separate sample. Instead of individual particles like classification approaches, these methods were trained to analyze sets of FIM images from a

sample. This approach could be used in fault detection applications to check if the particle population in a new batch of product is consistent with that typically generated by the manufacturing process.

This chapter provides an overview of the algorithms developed as part of the studies described in this thesis. Subsequent sections in this chapter contain a general description of each algorithm and a brief discussion of the strengths and weaknesses of each approach. Additional details about these algorithms can be found in later chapters in this thesis.

2.2 Method 1: KLD-MDS

2.2.1 Overview

Particles found in a single protein formulation are often highly heterogeneous in size and structure[8]. This variability often makes it difficult to compare particle populations between samples as the inherent variability in particle morphology within a sample can obfuscate any morphology differences between samples. A simple yet commonly used strategy to compare particle morphology between samples is to use graphical techniques to compare the distribution of common particle properties like size between samples. These methods use a small number of particle properties available from FIM instruments to represent the morphology of each particle and plot the distribution of either individual (i.e. as histograms) or sets of properties (i.e. as scatter plots) from each sample. The resulting plots can easily be compared between samples to determine if the samples exhibit different distributions of these properties. This approach is commonly used with size distributions in which histograms of particle diameters from different samples are compared against each other. A major weakness of these graphical

techniques is the limited number of particle properties that can be visualized and compared between samples as a time (i.e. up to three properties). To perform this analysis, the user must select which particle properties are to be used in the analysis--a decision that can bias the analysis so that only differences in the selected properties can be detected. Additionally, these graphical techniques are typically qualitative and cannot be used to directly quantify how different two samples are.

The initial method we developed to analyze FIM datasets was designed to perform a similar analysis to these graphical techniques. This initial algorithm used the Kullback-Liebler Divergence (KLD) to compare distributions of particle properties between samples quantitatively and without graphical techniques. To perform this analysis on FIM image sets from different samples, the user first decides which particle properties are to be used to compare samples. These particle properties are then computed for each particle in each FIM image set and the resulting dataset of particle property values are then used to estimate the probability density function (PDF) of these properties in the sample. These PDF estimates are constructed directly from particle properties of each FIM image using a kernel density estimate[9]. Property PDFs from two samples are then compared using a symmetrized KLD to quantify how different the particle populations in the two samples are from each other.

While the KLD measures how similar the particle morphology distributions in two samples are from each other, we were also interested in using these measurements to help compare particle morphologies between multiple samples at once. To perform this analysis on multiple samples, the KLD was first computed between all pairs of samples. The resulting pairwise similarity measurements were then analyzed via multidimensional scaling (MDS) to map each dataset onto points on a 2D plot that grouped samples with low KLD values and thus similar particle

morphology distributions together. More details on this KLD-MDS method can be found in the methods section in Chapter 3.

2.2.2 Application: Comparing particles made by accelerated stress conditions

In an early study[1], this algorithm was used to compare aggregates of a monoclonal antibody (mAb) donated by Medimmune, Inc. (Gaithersburg, Maryland) exposed to one of three accelerated stress conditions: freeze-thaw stress, low pH + shaking stress, and elevated temperature stress. The mAb was formulated in 230 mM KCl solution at pH 6 and in 2 mL plastic microcentrifuge tubes for each stress. Freeze-thaw stress consisted of exposing samples to 10 freeze-thaw cycles. Each cycle consisted of a 20 min freeze in a -80°C freezer followed by a 20 min thaw in a room temperature water bath. pH + shaking stress consisted of dialyzing mAb solution into a 20 mM citrate, 230 mM KCl solution at pH 3, dialyzing the solution into 20 mM histidine, 230 mM KCl solution at pH 6, and agitating the resulting protein solution on a plate shaker at 400 RPM overnight. Elevated temperature stress consisted of incubating samples at 60°C for three days. Three replicates were prepared per stress and imaged via a FlowCam VS system (Yokogawa Fluid Imaging Technologies, Inc., Scarborough, ME). Two separate FIM datasets were collected per replicate. Sample FIM images yielded by freeze-thaw and pH + shaking stress are shown in Figure 2.

The KLD-MDS algorithm described above was used to compare aggregates generated by each replicate of each stress. The PDF of particle aspect ratio was estimated for each sample using a kernel density estimate constructed from 1,000 FIM images per replicate and using a hard sphere kernel of radius 0.4. These PDFs were then analyzed using the KLD-MDS algorithm, yielding the 2D representations shown in Figure 3. The resulting embeddings generally group

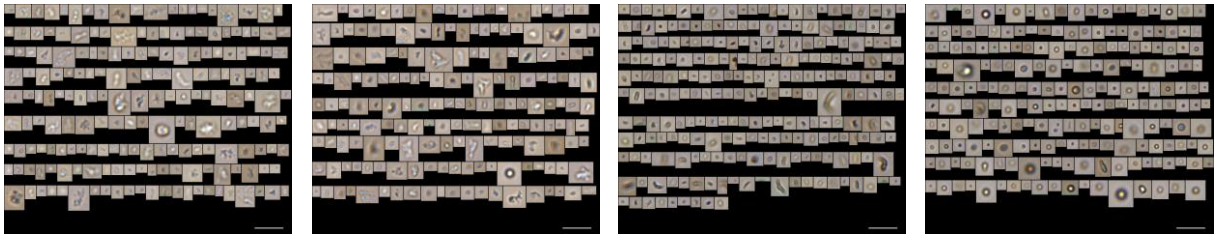

Figure 2: Sample FIM images of aggregates used in KLD-MDS/ConvNet algorithm development

FIM datasets by stress condition: FIM image sets from freeze thaw and pH samples appear clustered with other samples exposed to that stress in the right hand side of the figure while the elevated temperature samples appear on the left side of the figure. This plot indicates that aggregate populations generated by at least freeze-thaw and pH stress each exhibited a characteristic distribution of particle aspect ratios that can be used to distinguish particles made by that stress from those made by other stresses.

2.2.3 Benefits and Drawbacks

The KLD-MDS approach provides a quantitative way to compare particle populations between protein formulations—an analysis that previous machine learning methods for analyzing FIM images were not designed to do. Unlike previous graphical methods for doing this analysis, in principle this approach allows the user to easily analyze samples using a larger number of particle properties. Since this approach does not rely on plotting these properties, this analysis can be performed using many particle properties simultaneously assuming a sufficiently large number of FIM images are available to obtain an accurate PDF estimate. The MDS step also allows the user to compare these PDFs from multiple samples simultaneously which can be difficult to do using plots of the particle properties. While this approach still uses traditional

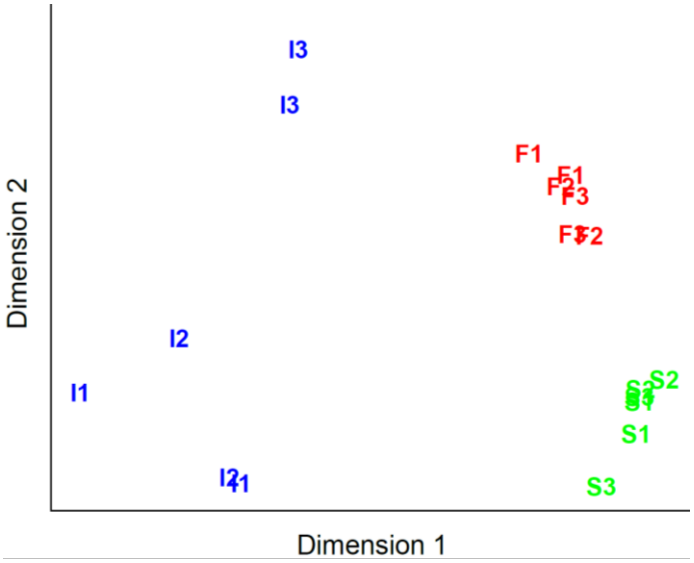

Figure 3: KLD-MDS plot of aggregates generated by accelerated stability stress
KLD-MDS plot of mAb aggregate populations generated by different stress conditions. Each alphanumeric label corresponds to the KLD-MDS representation a single FIM dataset generated by the stress denoted by the color letter (Red F for Freeze-thaw, Green S for pH and Shaking, Blue I for incubation at elevated temperature) and a replicate indicated by the number. Two technical replicates were taken via FIM per sample, resulting in two representations on this figure per sample. These representations were computed using particle aspect ratio. Dimensions 1 and 2 are the dimensions identified by MDS and do not have discernable physical meaning

particle properties to perform the analysis, this design ensures that the changes in particle morphology identified by this approach are related to changes in human-interpretable particle properties—behavior that is not guaranteed with the other approaches described in this thesis.

One of the most significant weaknesses of this approach is the need to estimate the PDF for a potentially large number of particle properties to perform the analysis. Obtaining accurate PDF estimates using kernel density estimates becomes increasingly difficult as the number of dimensions in the data increases. Due to the curse of dimensionality[9], the number of particles needed to obtain a PDF estimate at a fixed accuracy increases exponentially as the number of particle properties used in the analysis increases. While this approach can technically be used to analyze datasets using several particle properties, constructing accurate PDF estimates

using many particle properties would require the user to collect a massive number of FIM images. While >1,000 FIM images is sufficient to perform this analysis using a single particle property, a significantly larger image set (>10,000 images) must be collected to achieve similar algorithm performance using even a single additional particle property. This large particle requirement is not practical especially for samples with low particle concentrations like those encountered in commercialized protein formulations. This problem is exacerbated by the need to estimate the PDF for all samples in the analysis. This technique therefore requires a similarly large number of particles from all samples to be analyzed. Performing this analysis on a more appropriate number of particles requires the user to restrict the number of dimensions used to perform the analysis, resulting in many of the same issues described earlier with graphical techniques.

While KLD-MDS approach was useful for the applications we investigated initially, this flaw invalidated the logic behind the design of this algorithm. This flaw combined with other issues in this *ad hoc* approach ultimately led to us abandoning the algorithm in later projects. However, the early results obtained even with this crude algorithm suggested that the particle morphology information contained in FIM images was sensitive to the history of the sample and the types of stresses it was exposed to. The results obtained with this method in Chapter 3 also suggested that changes in particle morphology were correlated with different adverse responses in patients. While this technique was abandoned, this early success motivated us to iterate on this algorithm and develop effective methods for comparing particle morphology distributions.

2.3 Method 2: ConvNet Classifiers

2.3.1 Overview

A key limitation of many previous methods for analyzing particle morphology is the difficulty of developing a descriptive, usually low-dimensional representation of the particle morphology information encoded in FIM images. In both our first approach and other previous approaches this representation was constructed using sets of traditional particle properties returned by FIM instruments. Unfortunately, these particle properties contain limited particle morphology information and may not capture the most relevant particle morphology features to include in the analysis. It is desirable for FIM image analysis to develop more sophisticated image features and, since identifying and designing a quantitative measure of more relevant image features is time consuming, for this feature engineering to be performed automatically.

Convolutional neural networks (ConvNets) are an increasingly popular family of models designed to perform image analysis tasks in large part due to their ability to automatically learn image features from large image sets[10,11]. ConvNets primarily consist of a sequential series of convolutional layers, processing units that are designed to detect how well regions of an input image match small (usually 3 x 3) pixel patterns that the layer is designed to detect. These layers accept image data and return feature images that describe how well regions of the input match these patterns. The feature images returned by this layer can then be analyzed by an additional convolutional layer, allowing the subsequent layer to analyze the image in terms of the patterns identified by the previous layer. Stacking these layers on top of each other thus allows deeper layers of the network to detect complex pixel patterns in the raw input image. The final feature images are then processed using a few additional layer types to compute the final output the network is designed to yield such as an image classification.

Much of the usefulness of ConvNets stems from their ability to learn relevant image features for a given image analysis task directly from image data. The pixel patterns detected in each convolutional layer are optimized by training the network on images along with the desired network output. This feature allows us to leverage the large number of FIM images available in a typical FIM dataset to automatically "learn" relevant image features that are useful at mapping FIM images onto a desired output. These image features are usually only computed as an intermediate step in the network towards some other desired output and are not directly outputted or optimized. However, ConvNets can be trained to optimize and return these image features directly and an approach for doing so will be described later in this chapter. The ability of the network to automatically learn effective image features is incredibly useful for image analysis and has led to these networks being used to perform image analysis in areas ranging from traditional computer vision tasks (e.g. object[12,13] and facial[14] recognition) to image analysis tasks in science and medicine (e.g. cancer diagnosis[15–17]).

The first algorithm that was developed for ConvNet image analysis used these neural networks to perform image classification. To perform this analysis, a ConvNet is first trained on large numbers (e.g. >10,000) of FIM images of several different particle types labelled by the type of the particle in the image. During training, the network automatically learns image features that can be used to map these FIM images onto the label supplied with each image. Once trained, the ConvNet is then used to analyze new FIM images to predict which of the particle types the network was trained on best resembles the particle present in each inputted image.

2.3.2 Application: Identifying particles made by accelerated stress conditions

The first application investigated with this ConvNet method was identifying protein aggregates generated by different accelerated stress conditions[2]—conditions that we knew from earlier studies generated distinguishable aggregate morphologies. A ConvNet was trained to recognize FIM images of particles generated by one of four accelerated stress conditions: mAb aggregates generated by freeze-thaw and pH + shaking stress and intravenous immunoglobulin (IVIg; Gammagard Liquid; Baxter, Deerfield, IL) aggregates made by recirculation through one of two nominally identical piston fill-finish pumps. Freeze-thaw and pH aggregates were the same aggregates imaged and analyzed during development of the KLD-MDS algorithm. Pumping aggregates were generated by recirculating 45 mL of IVIg solution in 1x phosphate buffered saline (PBS) through one of two (denoted “Pump A” and “Pump B”) Filamatic FUS-10 pumps (Filamatic, Inc., Baltimore, Maryland) at a 200 mL/min flowrate for 9 minutes. The resulting aggregates were then imaged using a FlowCam VS. Sample FIM images used to train the algorithm are shown in Figure 2

50,000 labelled FIM images of each particle type were used to train a ConvNet to identify aggregates generated under each of these four classes. The trained network was then used to analyze an additional 10,000 images of each particle type that were excluded from algorithm training and without their corresponding label. Table 1 shows a confusion matrix for this neural network which shows the frequency that individual images of each particle type (Table 1, rows) were identified as each of the four particle types the network was trained to identify (Table 1, columns). The accuracy of this network on individual images (i.e. the average of the values along the diagonal of the table) was 78%—great accuracy given the heterogeneity of FIM images generated by a single condition and the similarity between images generated under different conditions. These results indicated that even individual particles within an FIM dataset

Table 1: ConvNet classifier accuracy on accelerated stability stress aggregates
A confusion matrix showing the fraction of individual FIM images from four samples (table rows) that were identified as each of the four possible classes the network was trained on (table columns). Values along the diagonal of the matrix indicate a correct classification

	Pump A	Pump B	pH + Shaking	Freeze-Thaw
Pump A	70%	30%	1%	0%
Pump B	18%	81%	1%	0%
pH + Shaking	0%	0%	71%	28%
Freeze-Thaw	0%	1%	8%	91%

can contain morphology information that could be used to identify the underlying conditions that sample was exposed to.

2.3.3 Benefits and Drawbacks

ConvNet-based methods in general benefit from their ability to automatically learn effective FIM image representations from the data. This feature not only allows the user to bypass the often-difficult process of selecting particle properties to use as image features in the analysis but can be used to easily include particle morphology features that are not normally accessible through simple particle properties returned by FIM instruments. These improved image features come at the cost of their interpretability. Unlike the human eye, ConvNets analyze images in 3x3-pixel segments rather than looking at the entire image at once. This difference in length scale allows the ConvNet to learn image features that are not necessarily human interpretable to differentiate between image types[18]. While these features will usually be effective at distinguishing between different particle morphologies for a given FIM image analysis task, it is difficult to interpret these features to identify the underlying structural changes in the FIM images these networks identify between particle types.

Another key benefit of ConvNets is the limited number of images necessary to analyze samples. While training these networks requires a substantial number of images (>10,000 per particle type), the trained network can easily be used to analyze individual FIM images from a sample. This contrasts with the previous method we developed which required at least 1,000 images from each sample to perform the analysis[1].

While ConvNets were found to be useful for analyzing FIM images, training these networks to perform classification was not useful for many protein aggregate analyses of interest in protein formulation development and production. This stems in part from the large number of particle-generating conditions that protein formulations may be exposed to in practice. Each of these conditions may generate a characteristic particle morphology that a classifier would need to correctly identify. However, classifiers do not effectively generalize beyond the image types they are trained to recognize and will therefore only be able to effectively analyze a subset of all possible particle types that may appear in a sample. Additionally, classifiers are inherently designed to only analyze individual images at a time and not particle populations. While this approach is useful in cases where analyzing the morphology of individual images is useful, for most protein analysis tasks we are generally more interested in techniques for comparing populations of these particles. However, some applications did require us to analyze the morphology of individual rather than sets of FIM images one of which is described in Chapter 8.

2.4 Method 3: ConvNets for Comparison

One strategy to mitigate some of the problems of the previous ConvNet approach is to set up the network to return image “embeddings” or a low-dimensional representation (i.e. set of image features) of the contents of the image instead of image classifications. These embeddings are

commonly used in facial recognition strategies[14,19] to represent images of different faces in a low-dimensional form that can easily be compared between images. These representations are much more generalizable than image classifications and can be used to analyze image types that were not explicitly included in the training set such as the faces of people not used to initially train a facial recognition algorithm. A ConvNet trained in this fashion on FIM image data would return similar low-dimensional representations that describe the morphology information contained in the image. Sets of these potentially very low-dimensional (e.g. 2D) embeddings can easily be analyzed using traditional statistical techniques such as hypothesis testing to determine if the sample exhibits a particle population consistent with that in another sample.

The final algorithm we developed used these facial recognition ideas combined with statistical techniques to determine if a sample exhibits a statistically different particle population than another sample. To perform this analysis, a ConvNet is first trained on labelled images of different particle types using a triplet loss approach borrowed from facial recognition[14]. This strategy involves training ConvNets to return two-dimensional embeddings of FIM images. These embeddings are optimized based on triplets of FIM images consisting of an image of a single particle type (anchor image), an FIM image of the same particle type (positive image), and an FIM image of a different particle type (negative image). The parameters of the network are optimized to reduce the distance between the embeddings of the anchor and positive images while increasing that between the embeddings of the anchor and negative image. This optimization results in a low-dimensional FIM image embedding scheme that groups together images of the same particle type and separate images of different particle types.

The trained network can be used to convert FIM images from two samples to low-dimensional embeddings that can easily be compared graphically. Additionally, goodness-of-fit hypothesis testing can be used to check if the FIM image embeddings from one sample are consistent with

those obtained from a second sample. To perform hypothesis testing, one of the samples to be compared is selected as a baseline sample with the other used as a test sample. The embeddings from the baseline sample are then used to construct a kernel density estimate of the PDF of embeddings from that sample. Goodness-of-fit hypothesis testing is then used to analyze small sets (c.a. 20-200) of FIM image embeddings from the test sample to test the null hypothesis that these embeddings are consistent with the embedding PDF estimated for the baseline sample. Rejection of this null hypothesis indicates that the two samples exhibited statistically significantly different particle morphology distributions. This hypothesis testing can be performed on multiple subsamples from the test sample and the rejection rate or the fraction of these subsamples that can be distinguished from the baseline sample is recorded as a measure of particle morphology distribution similarity. More details about this algorithm can be found in Chapter 4 and Appendix A while examples of how this algorithm can be used are found in Chapters 5-7.

This algorithm blends together features from the previous algorithms while also fixing a few problems with these earlier approaches. While this approach analyzes particle morphology distributions in a similar fashion as the initial KLD-MDS approach, this algorithm compares particle morphology distributions using hypothesis testing rather than the KLD metric. This change allows the user to easily determine if a difference in particle morphology distribution is statistically significant. This analysis also uses the potentially more descriptive but harder-to-interpret image features derived from ConvNets rather than combinations of simple particle properties. The low dimensionality of these ConvNet embeddings in addition to the use of hypothesis testing to compare samples also dramatically reduces the number of particles required to perform the analysis relative to the KLD-MDS algorithm: while >10,000 FIM images are required from a few samples to train the ConvNet and set up a baseline sample, very few (>20) images are required from test samples in order to compare them against a baseline

sample. Compared to the original ConvNet approach, this algorithm can more easily be used to analyze sets of FIM images and offers much better generalization to particle types that were not included in the ConvNet training.

One problem with triplet loss-derived embeddings is that, as a supervised technique (i.e. one that is trained on labelled data), the morphology information used to compute FIM embeddings is biased towards particle morphology features that discriminate between particles assigned different labels. Thus, if the particles assigned a single label exhibit multiple types of particle morphologies the algorithm does not reward an embedding scheme that maps these different particle types to different locations in the embedding space. This contrasts with unsupervised (i.e. label-free) techniques for computing similar embeddings such as variational autoencoders that, while not capable of extreme dimensionality reduction like the triplet loss algorithm, learn features that more heavily depend on the information in the image[20]. However, it may be difficult for these techniques to learn discriminative FIM image features since images from different samples often exhibit visually similar particles and backgrounds. This visual similarity may be especially problematic for features learned to minimize an image reconstruction error due to the very limited morphology information in individual pixels in FIM images. Additionally, as shown in subsequent chapters embeddings derived by triplet loss can still divide particle populations assigned a single label into subpopulations of visually similar aggregate types. While this behavior is not guaranteed by the loss function used to train these networks, this behavior does suggest that networks trained in this fashion can be used to capture particle morphology information beyond what is necessary distinguish between the labelled particle types.

It should also be noted that how the ConvNet is trained using this approach can influence the final performance of the model[21]. The triplets used to optimize the model's parameters are typically mined from all valid FIM image triplets during training due both to the large number of

valid triplets available and issues that can occur when training on triplets that have either very high or zero triplet loss. While this mining is critical to ensure the network trains effectively, how this mining is performed can impact the final FIM image embeddings the trained model returns and may influence the final conclusions drawn from that model. It is essential to compare the conclusions drawn from a ConvNet trained in this fashion against ConvNets trained using a simpler approach (e.g. classification) to ensure that the results do not depend on how the method was trained. However, the FIM datasets presented in this thesis were relatively robust against these changes in mining strategy and the results based on this algorithm were not significantly impacted by changes to the triplet mining strategy.

2.5 References

1. Maddux NR, Daniels AL, Randolph TW. Microflow Imaging Analyses Reflect Mechanisms of Aggregate Formation: Comparing Protein Particle Data Sets Using the Kullback-Leibler Divergence. *J Pharm Sci*. 2017;106(5):1239-1248. doi:10.1016/j.xphs.2017.01.030

2. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

3. Daniels AL, Calderon CP, Randolph TW. Machine Learning & Statistical Analyses for Extracting and Characterizing “Fingerprints” of Antibody Aggregation at Container Interfaces from Flow Microscopy Images. *Biotechnol Bioeng*. 2020.

4. Strehl R, Rombach-Riegraf V, Diez M, et al. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: A novel multiparametric image filter for sub-visible particles in microflow imaging analysis. *Pharm Res*. 2012;29(2):594-602. doi:10.1007/s11095-011-0590-7

5. Saggu M, Patel AR, Koulis T. A Random Forest Approach for Counting Silicone Oil Droplets and Protein Particles in Antibody Formulations Using Flow Microscopy. *Pharm Res*. 2017;34(2):479-491. doi:10.1007/s11095-016-2079-x

6. Kasimbeg PNO, Cheong FC, Ruffner DB, Blusewicz JM, Philips LA. Holographic Characterization of Protein Aggregates in the Presence of Silicone Oil and Surfactants. *J Pharm Sci*. 2018;108(1):155-161. doi:10.1016/j.xphs.2018.10.002

7. Probst C. Characterization of Protein Aggregates, Silicone Oil Droplets, and Protein-Silicone Interactions Using Imaging Flow Cytometry. *J Pharm Sci*. 2019:1-11. doi:10.1016/j.xphs.2019.05.018

8. Gambe-Gilbuena A, Shibano Y, Krayukhina E, Torisu T, Uchiyama S. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images. *J Pharm Sci*. 2020;109(1):614-623. doi:10.1016/j.xphs.2019.10.034

9. Scott DW. *Multivariate Density Estimation*.; 2015.

10. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444. doi:10.1038/nature14539

11. Goodfellow I, Bengio Y, Courville A. *Deep Learning*.; 2016. http://www.deeplearningbook.org/.

12. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. 2015:1-14.

13. Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions. *IEEE Conf Comput Vis Pattern Recognit*. 2015. doi:10.1109/CVPR.2015.7298594

14. Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2015;07-12-June:815-823. doi:10.1109/CVPR.2015.7298682

15. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118. doi:10.1038/nature21056

16. Arindra A, Setio A, Ciompi F, et al. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. *IEEE Trans Med Imaging*. 2016;35(5):1160-1169. doi:10.1109/TMI.2016.2536809

17. Kobayashi H, Lei C, Wu Y, et al. Intelligent whole-blood imaging flow cytometry for simple, rapid, and cost-effective drug-susceptibility testing of leukemia. *Lab Chip*. 2019;19(16):2688-2698. doi:10.1039/c8lc01370e

18. Wang H, Wu X, Huang Z, Xing EP. High Frequency Component Helps Explain the Generalization of Convolutional Neural Networks. 2019;(Remark 1):8684-8694. http://arxiv.org/abs/1905.13545.

19. Sun Y, Chen Y, Wang X, Tang X. Deep learning face representation by joint identification-verification. *Adv Neural Inf Process Syst*. 2014;3(January):1988-1996.

20. Kingma DP, Welling M. Auto-encoding variational bayes. *2nd Int Conf Learn Represent ICLR 2014 - Conf Track Proc*. 2014;(Ml):1-14.

21. Hermans A, Beyer L, Leibe B. In Defense of the Triplet Loss for Person Re-Identification. 2017. http://arxiv.org/abs/1703.07737.

Chapter 3: Flow microscopy imaging is sensitive to characteristics of subvisible particles in Peginesatide formulations associated with severe adverse reactions

Published as: Daniels AL, Randolph TW. Flow Microscopy Imaging Is Sensitive to Characteristics of Subvisible Particles in Peginesatide Formulations Associated With Severe Adverse Reactions. J Pharm Sci. 2018;107(5):1313-1321.

3.1 Abstract

The presence of subvisible particles in formulations of therapeutic proteins is a risk factor for adverse immune responses. Although the immunogenic potential of particulate contaminants likely depends on particle structural characteristics (e.g., composition, size, and shape), exact structure-immunogenicity relationships are unknown. Images recorded using flow imaging microscopy reflect information about particle morphology, but flow microscopy is typically used to determine only particle size distributions, neglecting information on particle morphological features that may be immunologically relevant. We recently developed computational techniques that utilize the Kullback-Leibler divergence and multidimensional scaling to compare the morphological properties of particles in sets of flow microscopy images. In the current work, we combined these techniques with expectation maximization cluster analyses, and used them to compare flow imaging microscopy datasets that had been collected by the US FDA after severe adverse drug reactions (including seven fatalities) were observed in patients that had been administered some lots of peginesatide formulations. Flow microscopy images of particle

populations found in the peginesatide lots associated with severe adverse reactions in patients were readily distinguishable from images of particles in lots where severe adverse reactions did not occur.

3.2 Introduction

Protein therapeutics offer numerous clinical benefits, and now comprise the fastest-growing class of drugs[1]. A challenge in the development of protein therapeutics is that they may elicit adverse drug reactions (ADRs) which include acute responses such as anaphylaxis during IV administration, or long-term adverse reactions such as immune responses wherein patients produce anti-drug antibodies (ADAs)[2–4]. The majority of current protein therapeutics are immunogenic in at least some patients[5], and in some cases (e.g., interferon beta[6–8]) adverse immune responses may be observed in up to half of patients treated, reducing efficacy[8]. Adverse immune responses can result in clinical trial failures[9].

There are a number of potential causes and risk factors associated with ADRs against protein therapeutics[10–17]. Among these risk factors is the presence of aggregates within protein formulations[11,18–28]. Numerous animal studies[29], human clinical studies[30–34] and *in vitro* studies[35–37] have associated particulate contaminants with infusion reactions, anaphylaxis, and activation of the innate and adaptive immune system[38,39]. Aggregation occurs as a result of various stresses to which proteins may be exposed, and different stresses such as freeze-thawing, exposure to air-water interfaces, pH extremes, elevated temperatures or chemical degradation produce different distributions of aggregates that are polydisperse in size and morphology[40]. *In vivo*, these aggregate populations may provoke different levels and types of immune responses[22,36,40]. For example, in one study, protein aggregates produced by process-related

conditions and low pH were not immunogenic[41]. Another study showed that larger, insoluble aggregates found in an antibody formulation after UV-light exposure were more immunogenic than soluble oligomeric aggregates of the same protein. At the present time, it is unclear which characteristic(s) of protein aggregates dictate their immunogenicity, in part because of the difficulties involved analyzing the different populations of particles (e.g., particles generated through different mechanisms of formation) that may be present in a given sample. Better techniques for characterizing aggregates are necessary in order to identify the features of protein aggregates that influence their ability to trigger ADRs upon administration—features that could be then monitored to assess the risk of ADRs and allow the most dangerous aggregate populations to be identified and prioritized for removal.

Protein drug manufacturers frequently use flow-imaging microscopy (FIM) to monitor the concentrations of micron-sized subvisible particles (e.g., protein aggregates, silicone oil droplets, air bubbles) present in protein formulations. In this technique, a sample is pumped through a microfluidic channel where a microscope records digital images of particles of size greater than about 2 µm. This technique yields either grayscale or color images representative of the 10^4- 10^6 individual particles larger than 2 µm typically present in a given sample. These image datasets are frequently large, with data file sizes of up to a gigabyte per sample. These collections of images potentially offer a wealth of particle structural information, but FIM is frequently used to obtain only particle size distributions as a histogram for a given sample. While convenient, this practice neglects other potentially relevant morphological features that could be extracted from these images. We hypothesize that the neglected information about particle morphology contained in the rich data sets generated by FIM could be relevant to determining risk of ADRs from particles within a protein formulation.

We recently developed a technique[42] to analyze collections of FIM datasets in order to differentiate between various populations of particles represented in the datasets. In this technique, the distributions of particle properties in each sample are compared to the distributions of properties in other samples via the symmetrized Kullback-Leibler divergence (KLD). A matrix of pairwise values of this divergence can then be processed via multidimensional scaling (MDS) to obtain a low-dimensional embedding of the data that captures the relative similarity between one dataset and the others included in the analysis. We previously demonstrated[42] that this technique can successfully differentiate between populations of particles formed in monoclonal antibody solutions that had been subjected to different aggregation-inducing stresses (freeze-thawing, shaking and pH changes, and elevated temperatures).

A recent study by the US FDA associated elevated levels of nano- and microparticles found in a marketed formulation of peginesatide (Omontys®; Affymax, Inc., Cupertino, CA) with severe ADRs in patients[43]. The drug, an erythropoiesis-stimulating agent consisting of a covalently dimerized synthetic peptide linked to polyethylene glycol, received FDA approval in 2012 for two formulations: a single-use vial formulation (SUV) and multi-use vial (MUV) formulation. The two formulations contain peginesatide at the same concentration, but have different excipients[44]. Although the SUV formulation was used predominantly during the clinical trials, only the MUV formulation was marketed. The marketed MUV formulation was linked to 49 cases of anaphylaxis (7 of which were fatal) and a hypersensitivity rate of 3.5 per 1000 exposed patients—significantly higher than the 0.84 per 1000 exposed patients rate that had been noted for the SUV formulation during the clinical trials. After the product was recalled voluntarily, the FDA investigated both the SUV and MUV formulations, conducting a variety of analyses in search of potential causes of the severe ADRs. Standard testing of the SUV and MUV formulations revealed that both formulations conformed to product specifications, including the

pharmacopeial limitations on the concentrations and size distributions for particles described by USP <788>. However, although both formulations met current limitations on particle content, it was discovered that the marketed MUV formulation had higher and more variable concentrations of subvisible particles than did the SUV formulation. Direct causality could not be established, but the analysis conducted by the FDA found that elevated subvisible particle content in the MUV formulations compared to that in the SUV formulations was associated with the observed increased hypersensitivity reactions seen for the marketed MUV formulation of peginesitide[43].

The FDA study[43] found that higher particle levels in MUV formulations of Omontys® were associated with increased rates of ADRs, but did not examine whether the formulations differed in particle characteristics other than concentration. Such differences might reflect different mechanisms by which the particles formed in the SUV and MUV formulations, which in turn could affect the propensity of the particles to generate ADRs. In the present study, we apply our KLD-MDS approach to the flow microscopy image datasets collected by the FDA in order to discern whether flow microscopy imaging can be used to differentiate between the ADR-associated particles found in MUV formulations of Omontys® and the particle populations that did not provoke adverse responses found in SUV formulations.

3.3 Materials and Methods

3.3.1 Materials

Intravenous immunoglobulin (IVIg; GAMMAGARD LIQUID) was obtained from Baxter International (Deerfield, IL). 1x phosphate buffered saline (PBS) containing 144 mg/mL

potassium phosphate monobasic, 795 mg/mL potassium phosphate dibasic, and 9000 mg/mL sodium chloride was obtained from Gibco (Waltham, MA). Hellmanex III was obtained from Hellma Analytics (Mullheim, Germany). All other salts and materials used in buffer preparation were reagent grade or higher.

3.3.2 Flow-Imaging Microscopy (FIM)

Flow-imaging microscopy datasets from the peginesatide investigation were provided by the FDA under a Freedom of Information Act (FOIA) request. In their investigation of the drug[43], the FDA analyzed samples from several SUV and MUV lots using a FlowCam VS1 system (Fluid Imaging Technologies, Inc., Scarborough, ME). The instrument used an 80-µm flow cell and a 10x objective. 450 µL of sample were analyzed in each measurement.

FIM datasets that we obtained from the FDA had been collected in three sets of FlowCam measurements referred to as “experiments” in the original study. We will use their nomenclature and denote these data collections as “Experiment A”, “Experiment B” and “Experiment C”. Due to limited sample volume, FIM settings were optimized over the course of data collection and thus each experiment used slightly different FIM settings. Both SUV and MUV samples were measured in each experiment. Experiment A contains 12 FIM datasets taken from 4 MUV lots and 4 datasets taken from a single SUV lot. Experiment B contains 11 datasets taken from 4 MUV lots and 8 datasets taken from 3 SUV lots. Experiment C contains 12 datasets taken from 4 MUV lots and 12 datasets taken from 3 SUV lots. Representative images taken from Experiment C for both formulations are shown in Figure 4.

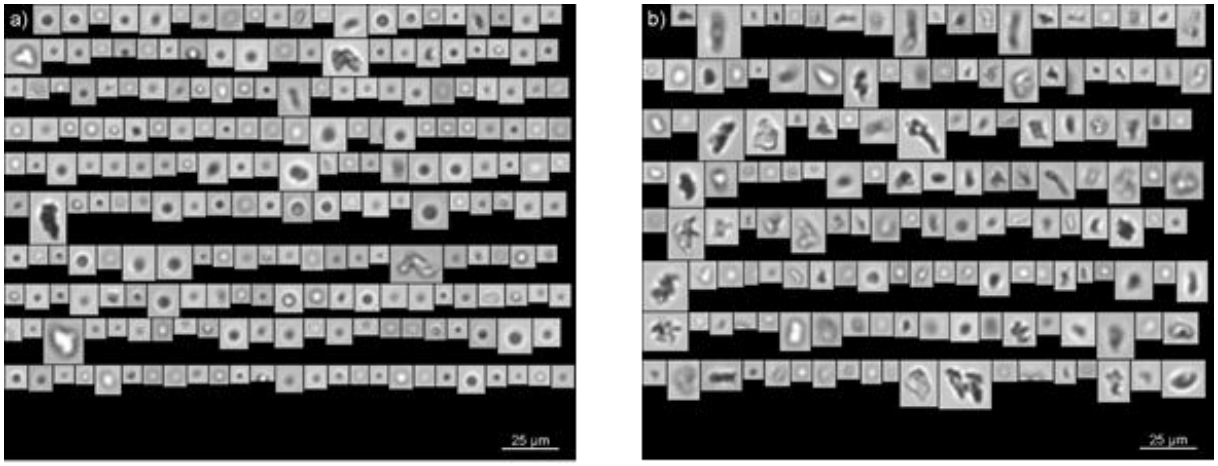

Figure 4: Sample peginesatide FIM images
Sample images taken from a) SUV and b) MUV samples analyzed in experiment C. Scale bars represent 20 µm

3.3.3 Image Analysis

Data analysis was performed in Python 3.6 (Python Software Foundation, OR). Images of the particles identified via the FlowCam instrument were imported into the software and segmented using custom image processing code to identify the particle-containing regions of the image. This analysis results in a “particle mask” or the portion of the raw image identified as a particle. The particle mask was then used to calculate several morphological properties for each particle. The area-based diameter of a particle was calculated by calculating the area of the particle mask in pixels and calculating the diameter of a circle with the same area. The aspect ratio of a particle was calculated by fitting the shape of the particle mask to an ellipse using principle components analysis (PCA) and dividing the length of the minor axis by the length of the major axis. The circularity of the particle was calculated by calculating the perimeter of the particle mask and dividing by the perimeter of a circle with the same area as the particle. The average particle intensity (i.e. grayscale color) was calculated by averaging the intensity of the pixels over the area of the particle. Histograms of the various particle properties were constructed in

order to visualize the distribution of properties within the datasets. The properties included in this analysis vary in scale. Aspect ratio, circularity, and average intensity values may vary from 0 to 1, but particle diameters can take on a much wider range of values. To avoid bias resulting from these differences in scale, each property was normalized by subtracting the mean value and dividing by the standard deviation before further analysis, effectively giving each observed property equal weighting in our analysis.

3.3.4 Kullback-Leibler Divergence (KLD)

The Kullback-Leibler Divergence is an information theory metric that can be used to estimate the similarity between two probability distribution functions (PDFs). A low value of this divergence indicates little difference in the distributions of particle properties and increasingly higher values indicate more strongly dissimilar distributions.

The KLD between two PDFs $P(\vec{d})$ and $Q(\vec{d})$ over a vector of particle properties $\vec{d}$ can be calculated by:

$$K(P||Q) = \int P(\vec{d}) \ln\left(\frac{P(\vec{d})}{Q(\vec{d})}\right) \mathrm{d}\vec{d} \tag{1}$$

where $K(P||Q)$ is the KLD between distributions $P(\vec{d})$ and $Q(\vec{d})$ and $\vec{d}$ is a vector containing normalized values of the four particle properties included in this analysis: diameter, aspect ratio, circularity, and average intensity. In this analysis $P(\vec{d})$ is an estimate of the distribution of particle properties in a FIM measurement containing $n_p = 2000$ particles with properties $\vec{d}_k$

where $k = 1 \ldots n_p$ indexes particles. $P(\vec{d})$ is estimated from $\vec{d}_k$ using a kernel density approximation with a hard sphere kernel of radius 0.6. $Q(\vec{d})$ is defined similarly for 2000 particles taken from a separate FIM dataset. As previously described[43], since $P(\vec{d})$ describes the frequency of specific values of $\vec{d}_k$ appearing in n_p particles, we can approximate $K(P||Q)$ as:

$$K(P||Q) = \frac{1}{n_p}\sum_{k=1}^{n_p} \ln\left(\frac{P(\vec{d}_k)}{Q(\vec{d}_k)}\right) \quad (2)$$

It is important to note that the KLD is not symmetric about the two distributions; $K(Q||P)$ will yield a different value than $K(P||Q)$ unless $P = Q$ due to the P weighting on the integral in eq. (1). This asymmetry can pose issues when trying to interpret the KLD as a measure of distribution similarity. We therefore use a symmetrized form of the KLD:

$$K(P,Q) = \frac{1}{2}\big((K(P||Q) + K(Q||P)\big) \quad (3)$$

where $K(P,Q)$ is the symmetrized KLD. Future mentions of the KLD will refer to this symmetrized form.

We wish to use eq. (2-3) to compare all possible pairs of FIM measurements in a dataset containing N FIM measurements. To compare dataset i to dataset j where $i = 1 \ldots N$ and $j = 1 \ldots N$ for both index datasets, we first compute P_i and Q_j, the distribution of particle properties in datasets i and j respectively. The particle properties that we utilize in this analysis are the particle diameter, aspect ratio, circularity, and average intensity, but other measures derived from the images, such as estimated masses of the individual particles[45] could also be used. We

can then use these distributions to construct a matrix A whose elements a_{ij} are the squared pairwise divergences between datasets i and j. a_{ij} can be calculated as:

$$a_{ij} = K\left(P_i, Q_j\right)^2 \tag{4}$$

3.3.5 Multidimensional Scaling (MDS)

While matrix A contains significant information about the similarity between pairs of datasets, it is challenging to extract global similarity information from these pairwise divergence values. We can use multidimensional scaling (MDS) to find a low-dimensional embedding of the datasets that best captures the pairwise similarity information contained in A as was described and performed in previous work[42]. MDS techniques are designed to operate on a matrix of pairwise distances like the distances between cities on a map. Although the KLD is a divergence between distributions and not a distance metric, in this analysis we will treat the symmetrized KLD as a distance measurement and use MDS techniques as a method to visualize the relative similarity between datasets.

In classical multidimensional scaling (CMDS), we seek values of hypothetical points $\vec{x}_i$ where $i = 1 \ldots N$ again indexes datasets whose values are set so that the distance between $\vec{x}_i$ and $\vec{x}_j$ (where $j = 1 \ldots N$ again indexes datasets) is similar in value to the corresponding entry in matrix A a_{ij}. To perform this analysis, we first apply double centering to A using the centering matrix J whose elements j_{kl} are defined as:

$$j_{kl} = \delta_{kl} - \left(\frac{1}{N}\right) \tag{5}$$

where $k = 1 \ldots N$ and $l = 1 \ldots N$ are now indices in the centering matrix and δ_{kl} is the Kronecker delta. This matrix can be used to double center matrix A or to subtract the mean of each row and each column from the matrix:

$$B = -\frac{1}{2}JAJ \tag{6}$$

where B is the double centered matrix. The minimization to find points $\vec{x}_i$ can now be written in terms of entries in the double centered matrix b_{ij} as:

$$\min_x \sum_{i=1}^{N}\sum_{j=1}^{N} \left\| b_{ij} - \left\|\overrightarrow{x_i} - \overrightarrow{x_j}\right\|^2 \right\|^2 \tag{7}$$

The values $\vec{x}_i$ that minimize equation (N) can be found by performing eigendecomposition analysis on matrix B. Matrix B can be represented in terms of a matrix of N eigenvalues and N corresponding eigenvectors:

$$B = Q\Lambda \mathrm{Q}^{-1} \tag{8}$$

where Λ is a diagonal matrix of eigenvalues and Q is a matrix of the corresponding eigenvectors. To find values two dimensional coordinates that satisfy the minimization in eq. (7), we simply select the two eigenvectors with the largest corresponding eigenvalues. $\vec{x}_i$ can then be calculated using:

$$\overrightarrow{x_i} = Q_2\Lambda_2^{1/2} \tag{9}$$

Where Λ_2 is a diagonal matrix of the two highest eigenvalues and Q_2 a matrix containing the corresponding eigenvectors. This equation yields a two-dimensional representation of the pairwise KLD values that can easily be plotted and analyzed.

3.3.6 KLD-MDS Plots

To compare the FIM datasets, the distribution of particle diameter, aspect ratio, circularity, and average intensity in each dataset were compared to those of the other datasets via the Kullback-Leibler divergence. The pairwise divergences were then analyzed via classical multidimensional scaling to obtain a 2D embedding for each dataset. The resulting 2D coordinates reflect the underlying similarity between datasets as measured by the KLD: two datasets that are similar as indicated by a low value of the KLD will generally appear closer together on the figure than two datasets that are more dissimilar.

It is important to note that MDS assigns coordinates to the datasets so that the distance between any two points describes the relative value of the corresponding value of the KLD. The distances between points on the 2D projection thus are the main quantities that can be used to interpret the FIM datasets described by the figure. The axes on which the coordinates are oriented are chosen to represent these distances and do not have any easily discernable physical meaning in terms of either the underlying particle properties or the values of the KLD between datasets. Since only the relative locations of the points in the embedding is significant, for ease in visualization we chose to rotate the axes obtained from KLD-MDS so that the average coordinate of both SUV and MUV datasets lay on the x-axis and that the SUV was on the left of the embedding.

KLD-MDS embeddings were obtained using data from Experiments A, B, and C individually. Additionally, we also performed the analysis using data from all three experiments at once. Since the FlowCam settings varied between the three experiments, it is possible that the differences in settings could introduce artifacts into the final KLD-MDS embedding. To account for these potential differences, data in each experiment was normalized to the mean and standard deviation of properties within that experiment before normalizing to the overall mean and standard deviation.

3.3.7 Clustering Analysis

KLD-MDS can be used to compare the types of particles present in the two formulations to determine whether the formulation influences the morphology of the particles present in a dataset. Since only a finite number of particles are present in a given sample, we expect some variation in the types of particles present between every pair of FIM measurements in a given dataset—even for samples taken from the same formulation. If the formulation significantly influences the morphology of the particles in the sample, two FIM measurements taken from different formulations will have more dramatic differences in particle morphology than two measurements taken from the same formulation. These differences in particle morphology will then be apparent as larger KLD values for pairs of measurements taken from different formulations than for measurements taken from the same formulation. This pattern in KLD values will manifest as clustering in the final embedding: FIM measurements on SUV samples will embed closer to other measurements performed on SUV samples and further away from MUV measurements. We can therefore identify a formulation-dependent shift in subvisible particle properties by assessing the clustering in the final KLD-MDS plots. If the samples can be

successfully clustered by formulation, then the subvisible particles likely exhibit a formulation-dependent shift in particle properties indicative of a change in the subvisible particle population.

The clustering in the KLD-MDS plots can be assessed visually to determine if the datasets cluster by formulation. Alternatively, various metrics can be used to quantitatively measure the resolution of the clustering. One such metric is the silhouette coefficient[46], which is a measure of how similar a given data point in a cluster is to other points in its cluster relative to its similarity to points outside its cluster. The silhouette coefficient for a given point x_i in a set of two clusters $si(x_i)$ is calculated as

$$si(x_i) = \frac{d_{inside}(x_i) - d_{outside}(x_i)}{max(d_{inside}(x_i), d_{outside}(x_i))} \quad (10)$$

Where $d_{inside}(x_i)$ and $d_{outside}(x_i)$ are the average Euclidean distance between point x_i and points inside the cluster x_i belongs to and outside the cluster, respectively. The silhouette coefficient can take on values from -1 to 1 with higher values indicating increasing relative similarity to the other points in the cluster x_i is assigned to. Since every point in the KLD-MDS dataset will have a silhouette coefficient associated with it, we report the average silhouette coefficient for SUV and MUV datasets in each of the 3 experiments.

We can also identify if the known clustering in formulation can be recovered using an unsupervised clustering technique. To accomplish this, expectation maximization[47] (EM) was applied to the unlabeled coordinates from each experiment to estimate two 2D Gaussian Distributions to represent the two expected clusters. Starting from an initial guess of the two distributions, EM iteratively calculates the probability of each point in the KLD-MDS embedding belonging to the two clusters and, using these probabilities as weights, re-estimating the mean

and covariance matrix of the two distributions. This process is repeated until two distributions that represent the clustering in the coordinates are obtained. Points in the KLD-MDS embedding with $p > 0.01$ for a given distribution were then assigned to the cluster that distribution represents. We anticipate that this process should group datasets by the formulation of the sample if the subvisible particles in the two formulations are significantly different in morphology.

3.3.8 Effect of Formulation Differences

The two approved Omontys® formulations contain identical concentrations of peginesatide as well as sorbitol, but the SUV uses a phosphate buffer and contains micromolar concentrations of Tween 20, whereas the MUV uses a methionine buffer and contains phenol as a preservative. These differences in formulation might have affected particle formation mechanism(s) and consequently the resulting particle morphologies in the two formulations. However, it is also possible that these differences might have affected the capability of FIM to accurately detect and measure morphological features of particles. For example, differences in refractive index (RI) can alter the apparent transparency of particles in FIM[48]. It is therefore possible that different populations of particles detected using KLD-MDS analysis could reflect formulation differences, rather than differences in particle morphologies.

To test whether the KLD-MDS analysis was detecting only formulation differences between SUV and MUV formulations rather than differences in the morphologies of particles within the respective formulations, we created a standard population of protein aggregates, and then spiked small amounts of these pre-formed aggregates into solutions whose excipient concentrations matched those of the SUV and MUV formulations. FIM and KLD-MDS analyses

were then conducted to determine whether the standard protein aggregate particles appeared to be different in the two formulations.

Preparation of suspensions of standard protein particles in MUV and SUV formulations: To generate a standard suspension of protein aggregates, 0.5 mg/mL IVIg in 1xPBS were centrifuged at 20,000 g at 4°C for 20 minutes to remove small aggregates. 0.9 mL aliquots of the supernatant were placed in 1.5 mL microcentrifuge tubes and exposed to six freeze-thaw cycles. Each cycle consisted of suspension in liquid nitrogen for 2 minutes followed by suspension in a hot water bath at 30 ºC for 6 minutes. Two buffers were made which matched the excipient profile of the SUV and MUV formulation as described in the prescribing information. The SUV formulation contained 47 mg/mL sorbitol, 2.3 mg/mL anhydrous sodium phosphate monobasic, 0.12 mg/mL sodium phosphate dibasic, and 0.04 mg/mL polysorbate 20. The MUV formulation contained 47 mg/mL sorbitol, 5 mg/mL phenol, 1.5 mg/mL L-methionine, and 0.6 mg/mL glacial acetic acid. The pH of the SUV and MUV formulations were 6.0 and 5.4, respectively. 50 µL aliquots of the standard suspension fo protein aggregates were mixed with 950 µL aliquots of either the SUV and MUV formulation buffers to form triplicate 1 mL aliquots of IVIg particles in the SUV and MUV formulations.

Analysis of standard protein particles spiked into MUV and SUV solutions: Flow-imaging was performed on three 300 µL aliquots of each sample using a FlowCam® VS (Fluid Imaging Technologies, Inc., Scarborough, ME) instrument using a 100-µm flow cell and the 10x objective. The flowcell was flushed with 1% Hellmanex III solution followed by water before running samples and with water between individual measurements. Images obtained from the FlowCam were analyzed using the KLD-MDS algorithm using the same particle properties used to analyze the full Omontys dataset (i.e. size, aspect ratio, circularity, and average intensity). Since the FlowCam model used in this analysis collects RGB color images as opposed to the

grayscale images available in the initial dataset collected by the FDA researchers, color images were converted to grayscale images prior to analysis. RGB pixel values were converted to grayscale intensities using the luminosity conversion:

$$c_{gray} = 0.2126c_{red} \rightleftarrows +0.7152c_{green} + 0.0722c_{blue} \quad (11)$$

where c_{gray} is a grayscale pixel value (i.e. pixel intensity) and c_{red}, c_{green}, and c_{blue} are the red, green, and blue channel values, respectively, of the corresponding RGB pixel.

3.4 Results

3.4.1 Image Analysis

Figure 5 (a, b) shows representative histograms of the particle size distributions obtained from the image analysis for randomly chosen SUV and MUV vials, respectively, taken from experiment C. The two samples have relatively similar size distributions and would otherwise be difficult to differentiate by visual analysis, especially in a quantitative manner.

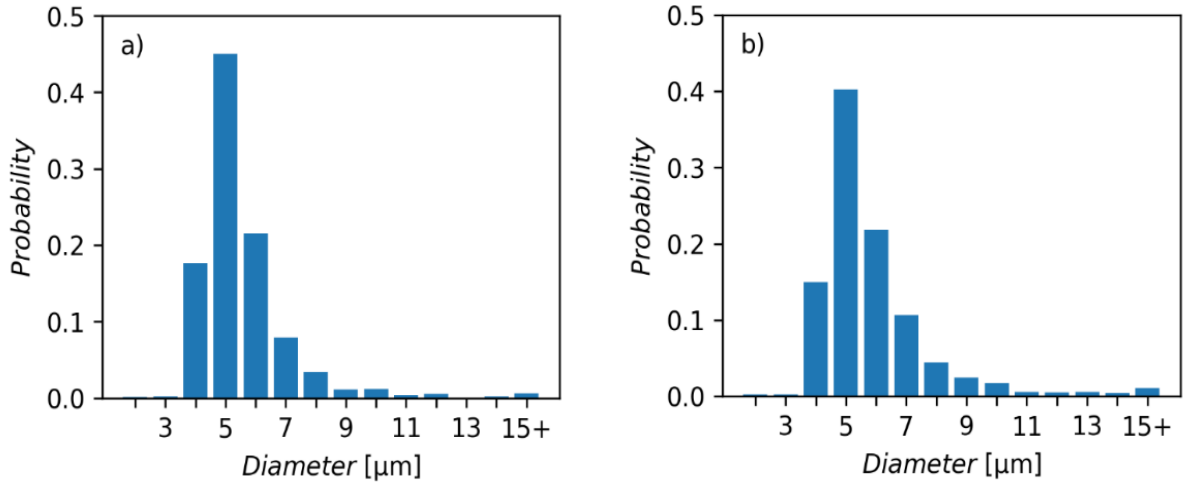

Figure 5: Peginesatide size distributions
Histograms of SUV and MUV particle sizes. Histograms of particle size of a random a) SUV and b) MUV sample taken from experiment C. The MUV sample shows the presence of generally larger particles than the SUV sample. However, the histograms otherwise indicate generally similar distributions of particle sizes.

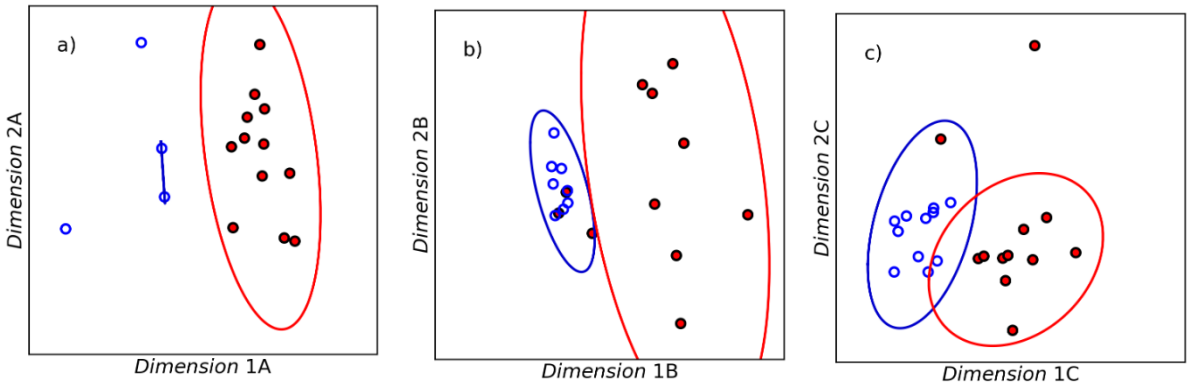

Figure 6: Peginesatide KLD-MDS plots (individual experiments)
KLD-MDS plots of the FIM data collected in (a) experiment A, (b) experiment B, and (c) experiment C. Plots were constructed using the distribution of particle diameter, aspect ratio, circularity, and average intensity in the FIM datasets. Also shown on these figures are the $p > 0.01$ regions of the 2D Gaussians obtained from expectation maximization as represented by the red and blue ovals. As can be seen visually, datasets in experiment B exhibit moderate clustering by formulation and datasets in experiments A and C exhibit much more substantial clustering by formulation. This clustering is confirmed by EM; most SUV datasets (open blue circles) are within the $p > 0.01$ region of the SUV cluster (blue oval) and most MUV datasets (red-filled circles) are within the $p > 0.01$ region of the MUV cluster (red oval).

3.4.2 KLD-MDS Plots

Figure 6 (a, b, c) shows the two-dimensional embeddings obtained from the KLD-MDS analysis for experiments A, B, and C, respectively. Grouping of samples by formulation can be observed in all three figures. For all three experiments SUV samples appear closer to other SUV samples than MUV samples in the KLD-MDS embedding. Although three MUV samples overlap with SUV samples in figure 6(b) and two samples are separated from the MUV samples in figure 3(c), in general MUV samples also generally exhibit clustering behavior.

3.4.3 Clustering Analysis

Table 2 shows the average silhouette coefficient for each formulation in each of the three experiments as well as the average overall silhouette coefficient for the experiment. Experiments A and C have approximately similar values of the silhouette coefficient while Experiment B has a significantly lower average silhouette coefficient as was expected from Figure 6(b). Note that, with the exception of MUV vials in experiment B, all formulations have moderate positive average silhouette coefficients, indicating that the datasets are generally clustered by formulation.

Table 2: Silhouette coefficients for peginesatide KLD-MDS plots

Experiment	SUV Silhouette Coefficient	MUV Silhouette Coefficient	Average Silhouette Coefficient
A	0.37	0.53	0.49
B	0.78	-0.04	0.31
C	0.71	0.25	0.48

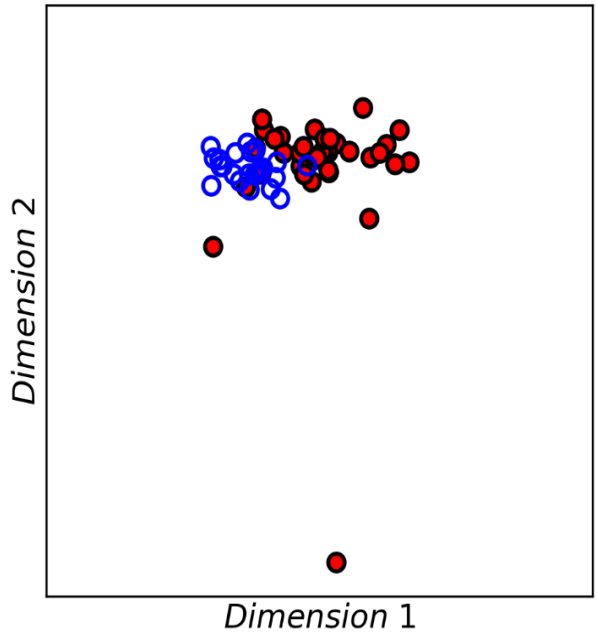

Figure 7: Peginesatide KLD-MDS plots (all experiments)
KLD-MDS plot showing data for all three experiments. Plots were constructed using the distribution of particle diameter, aspect ratio, circularity, and average intensity in the FIM datasets. Although the plot lacks resolved clusters for the SUV (open blue circles) and MUV samples (red-filled circles), the datasets still exhibit noticeable segregation by formulation indicative of a formulation dependency in the subvisible particle populations.

Figure 6(a, b, c) also shows the boundary of clusters obtained from EM analysis for Experiments A, B, and C, respectively. The boundaries represent the $p > 0.01$ probability region of the two 2D Gaussians representing the two clusters. The clusters obtained from EM generally separate the datasets by formulation. As shown in figure 6(a), 14 of the 16 datasets in Experiment A are correctly sorted by formulation. Similarly, 16 of the 19 datasets are correctly sorted by formulation for Experiment B and 21 of the 24 datasets are correctly sorted for Experiment C.

Figure 7 shows the two-dimensional embedding obtained from performing the KLD-MDS analysis using data from all three experiments at once. In spite of the differences in FlowCam analysis parameters used by the FDA for the three experiments, with proper normalization the

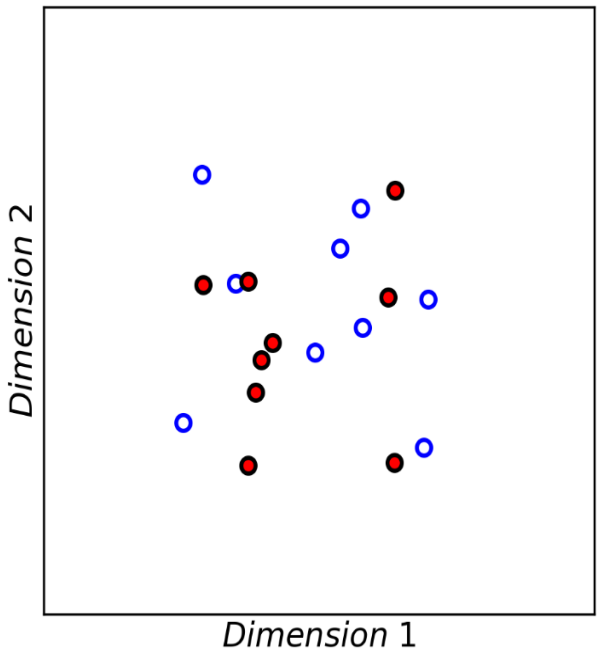

Figure 8: Formulation effect on KLD-MDS plots
KLD-MDS plot of the IVIg particles suspended in SUV (red-filled circles) and MUV (open blue circles) formulation buffers. As in figure 6 plots were constructed using the distribution of particle diameter, aspect ratio, circularity, and average intensity in the FIM datasets. Visually the clustering present in figures 6-7 is mostly absent when both formulations contain identical particles, suggesting that the refractive indices of the two formulations does not significantly influence the KLD-MDS analysis

KLD-MDS analysis still indicates a significant formulation dependency on the properties of subvisible particles present in the sample. Although only a single cluster appears in the figure, the coordinates are segregated by formulation within this cluster with the majority of SUV datasets appearing on the left of the cluster and the majority of MUV datasets appearing on the right.

3.4.4 Effect of Omontys formulations on FIM analysis

Figure 8 shows the results of KLD-MDS analysis of suspensions of standard protein aggregates in the MUV and SUV formulations. The populations of standard protein aggregates in the two

formulations were indistinguishable by KLD-MDS analysis, with silhouette coefficients of 0.023 and 0.14 for the SUV and MUV formulations, respectively.

3.5 Discussion

Automated image processing techniques are continuously growing in prominence in performance in tasks ranging from object recognition[49] to developing self-driving cars[50]. These image processing techniques have also shown potential in areas of biomedical research such as diagnosing lung cancer from CT imaging[51] and diagnosing melanoma from smartphone images[52]. In these cases, image analysis techniques could diagnose cancers with accuracy meeting or exceeding that of trained clinicians. These image processing techniques may greatly enhance the amount of subvisible particle information that can be extracted from FIM. Although machine learning techniques have recently been used to analyze these images, these approaches have primarily been used to perform simple classification tasks such as the differentiation of silicone oil droplets from protein aggregates[53–55]. It is anticipated that more sophisticated image analysis techniques such as this Kullback-Leibler divergence-based approach may allow users to perform more complicated characterization tasks.

KLD-MDS can be applied to FIM datasets to identify a difference in particle morphology between the two formulations of peginesatide. In the 2D embedding obtained from KLD-MDS,most FIM measurements in the three experiments cluster with other measurements taken from the same formulation, e.g., FIM datasets from MUV samples are closest to datasets from other MUV samples, and similarly FIM datasets from SUV samples are closest to datasets from other SUV samples. The clustering is most apparent for datasets collected in Experiments A and C. Although the measurements taken in Experiment B exhibit less well-resolved clusters,

the majority of datasets is still reasonably clustered in the KLD-MDS embedding. These formulation-dependent differences can also be identified when comparing FIM datasets from multiple experiments at once with proper normalization. Although the KLD-MDS embedding obtained from performing the analysis on the combined FIM data sets from all three experiments A, B and C lacks the more resolved clustering observed for the individual experiments, the new datasets still appear to be readily sorted by the formulation from which the dataset was obtained.

Clustering analysis by EM analysis was performed using a fixed number of clusters (two), as opposed to traditional clustering problems where the number of clusters may be a fitted variable. The clustering of these datasets as determined by application of EM algorithms agrees well with a simple visual analysis of the KLD-MDS plots, in which it is apparent that the data are largely segregated into two groups composed mostly of either SUV- or MUV-derived samples. Because the goal of this analysis is to see how well an unsupervised clustering technique can recover the anticipated formulation-dependent subvisible particle differences, fixing the number of clusters at two is sufficient for our purposes.

In the KLD-MDS analysis, the sets of data from the SUV and the MUV lots were relatively well-separated, so the value of p used to specify the decision boundary of the clusters calculated by EM had relatively little effect on which points were contained within the given clusters. However, we anticipate that the choice of *p* value will be more important for eventual applications of these techniques in process monitoring and control. For instance, consider a process-monitoring application in which FIM datasets recorded on new lots of a product are compared to sets of FIM measurements from older lots that are known to meet product specifications. p is the significance level threshold that determines whether or not a given lot has particles that match those found in the lots that meet product specifications. In this example p should be set high

enough to identify samples with particles that dramatically vary in subvisible particle populations and potential immunogenicity from the normal product, but low enough to avoid unnecessary process downtime due to an incorrectly identified process upset. Optimization of the value of p needed to balance these two risks was outside the scope of this study.

The clustering of datasets by formulation suggests that populations of subvisible particles in the two peginesatide formulations exhibit significantly different morphologies; these differences in particle characteristics are larger than any differences between samples of the same formulation. These subtle differences in particle morphology are difficult to detect, even at a qualitative level, using currently standard FIM analyses like the particle size distribution histograms shown in Figure 5. Although histograms such as those shown in Figure 5 are visually difficult to differentiate even in a qualitative manner, the KLD-based approach is capable of identifying a quantitative difference in subvisible particle populations between the two formulations. This variation correlates with the difference in the frequency of severe ADRs that was observed in the clinic, suggesting that a change in particle morphology could have contributed to the change in immunogenicity.

The SUV and MUV formulations of peginesatide were slightly different in composition and were filled into different containers. These differences in formulation and container-closure systems likely contributed to differences in the mechanisms by which the particles were created, and in turn resulted in differences in particle morphology that could be detected by our KLD-MDS analyses of FIM datasets. An alternative explanation is that apparent differences in particle populations detected in SUV and MUV datasets are the result of formulation-generated biases in the FIM analyses, rather than actual differences in particle morphology. This alternative explanation can be discounted, because an analysis conducted on suspensions of standard

aggregates in the same formulations yielded no detectable differences in particle populations that could be ascribed to formulation effects on FIM images.

Although the FDA researchers had previously identified differences in the numbers of subvisible particles found in the clinical and marketed Omontys® formulations[43], we stress that we have identified a separate difference in between their populations of subvisible particles. The FDA reported that the MUV formulation generally had higher concentrations of subvisible particles than did the SUV formulation. In contrast, our analysis is not influenced by the concentration of particles in the sample but is instead focused on identifying differences in morphology between *populations* of particles in the two samples. Our analysis therefore identifies particle morphology as a separate factor that could have contributed to the ADRs to the marketed Omontys® formulation.

Our findings in conjunction with the earlier findings of the FDA[43] indicate that the two Omontys® formulations exhibited substantial differences in subvisible particle populations. Although both formulations met particle concentration limitations set by USP <788> for particles of size larger than 10µm and larger than 25 µm[43], subvisible particle populations for the two formulations differed both in particle concentration and, as is evident from our KLD-MDS analysis, particle morphological properties. Although they are associated with the serious ADRs experienced by patients receiving the MUV formulations of Omontys®, neither the population distributions of subvisible particles nor their respective morphologies can be causally linked to the ADRs. It is apparent in retrospect, however, that had the differences in concentrations and morphologies of subvisible particles between the safe, clinically-tested SUV formulation and the ADR-provoking MUV formulation been known, a red flag should have been raised prior to initiation of marketing.

3.6 Acknowledgements

This study was funded by the National Institutes of Health, grant RO1 EB006006.

3.7 References

1. Walsh G. Biopharmaceutical benchmarks 2018. Nat Biotechnol. 2018;36(12):1136-1145. doi:10.1038/nbt.4305

2. Schellekens H. Immunogenicity of therapeutic proteins: Clinical implications and future prospects. Clin Ther. 2002;24(11):1720-1740. doi:10.1016/S0149-2918(02)80075-3

3. Schellekens H. Immunogenicity of therapeutic proteins. *Nephrol Dial Transplant*. 2003;18(7):1257-1259. doi:10.1093/ndt/gfg164

4. Tamilvanan S, Raja NL, Sa B, Basu SK. Clinical concerns of immunogenicity produced at cellular levels by biopharmaceuticals following their parenteral administration into human body. *J Drug Target*. 2010;18(7):489-498. doi:10.3109/10611861003649746

5. Singh SK. Impact of product-related factors on immunogenicity of biotherapeutics. *J Pharm Sci*. 2011;100(2):354-387.

6. Hartung HP, Munschauer F, Schellekens H. Significance of neutralizing antibodies to interferon beta during treatment of multiple sclerosis: Expert opinions based on the Proceedings of an International Consensus Conference. *Eur J Neurol*. 2005;12(8):588-601. doi:10.1111/j.1468-1331.2005.01104.x

7. Godin DS, Frohman EM, Hurwitz B, et al. Neutralizing antibodies to interferon beta: Assesment of their clinical and radiographic impact: An evidence report - report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology*. 2007;68(13):977-984.

8. Malucchi S, Sala A, Gilli F, et al. Neutralizing antibodies reduce the efficacy of βIFN during treatment of multiple sclerosis. *Neurology*. 2004;62(11):2031-2037. doi:10.1212/01.WNL.0000129265.73259.9E

9. Ridker PM, Tardif J-C, Amarenco P, et al. Lipid-Reduction Variability and Antidrug-Antibody Formation with Bococizumab. *N Engl J Med*. 2017;376(16):1517-1526. doi:10.1056/nejmoa1614062

10. Schellekens H. Bioequivalence and the immunogenicity of biopharmaceuticals. *Nat Rev Drug Discov*. 2002;1(6):457-462. doi:10.1038/nrd818

11. Hermeling S, Crommelin DJA, Schellekens H, Jiskoot W. Structure-immunogenicity relationships of therapeutic proteins. *Pharm Res*. 2004;21(6):897-903. doi:10.1023/B:PHAM.0000029275.41323.a6

12. Sauerborn M, Brinks V, Jiskoot W, Schellekens H. Immunological mechanism underlying the immune response to recombinant human protein therapeutics. *Trends Pharmacol Sci*. 2010;31(2):53-59. doi:10.1016/j.tips.2009.11.001

13. van Beers MMC, Jiskoot W, Schellekens H. On the role of aggregates in the immunogenicity of recombinant human interferon beta in patients with multiple sclerosis. *J Interf cytokine Res*. 2010;30(10):767-775.

14. Van Beers MMC, Bardor M. Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. *Biotechnol J*. 2012;7(12):1473-1484. doi:10.1002/biot.201200065

15. Torosantucci R, Schoneich C, Jiskoot W. Oxidation of therapeutic proteins and peprides: structural and biological consequences. *Pharm Res*. 2014;11(2):99-109.

16. Ratanji KD, Derrick JP, Dearman RJ, Kimber I. Immunogenicity of therapeutic proteins: Influence of aggregation. *J Immunotoxicol*. 2014;11(2):99-109. doi:10.3109/1547691X.2013.821564

17. Wang W, Singh SK, Li N, Toler MR, King KR, Nema S. Immunogenicity of protein aggregates - Concerns and realities. *Int J Pharm*. 2012;431(1-2):1-11. doi:10.1016/j.ijpharm.2012.04.040

18. Chisholm CF, Baker AE, Soucie KR, Torres RM, Carpenter JF, Randolph TW. Silicone Oil Microdroplets Can Induce Antibody Responses Against Recombinant Murine Growth Hormone in Mice. *J Pharm Sci*. 2016;105(5):1623-1632. doi:10.1016/j.xphs.2016.02.019

19. Den Engelsman J, Garidel P, Smulders R, et al. Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. *Pharm Res*. 2011;28(4):920-933. doi:10.1007/s11095-010-0297-1

20. Rosenberg AS. Effects of protein aggregates: An immunologic perspective. *AAPS J*. 2006;8(3):E501-E507. doi:10.1208/aapsj080359

21. Chisholm CF, Soucie KR, Song JS, et al. Immunogenicity of Structurally Perturbed Hen Egg Lysozyme Adsorbed to Silicone Oil Microdroplets in Wild-Type and Transgenic Mouse Models. *J Pharm Sci*. 2017;106(6):1519-1527.

22. Fradkin AH, Carpenter JF, Randolph TW. Immunogenicity of aggregates of recombinant human growth hormone in mouse models. *J Pharm Sci*. 2009;98(9):3247-3264. doi:10.1002/jps.21834

23. Fradkin AH, Carpenter JF, Randolph TW. Glass Particles as an Adjuvant: A Model for Adverse Immunogenicity of Therapeutic Proteins. *J Pharm Sci*. 2011;100(11):4953-4964.

24. Van Beers MMC, Gilli F, Schellekens H, Randolph TW, Jiskoot W. Immunogenicity of recombinant human interferon beta interacting with particles of glass, metal, and

polystyrene. *J Pharm Sci*. 2012;101(1):187-199. doi:10.1002/jps.22744

25. Shomali M, Freitag A, Engert J, et al. Antibody responses in mice to particles formed from adsorption of a murine monoclonal antibody onto glass microparticles. *J Pharm Sci*. 2014;103(1):78-89. doi:10.1002/jps.23772

26. Freitag AJ, Shomali M, Michalakis S, et al. Investigation of the immunogenicity of different types of aggregates of a murine monoclonal antibody in mice. *Pharm Res*. 2015;32(2):430-444. doi:10.1007/s11095-014-1472-6

27. Shomali M, Tanriverdi S, Freitag AJ, et al. Dose levels in particulate-containing formulations impact anti-drug antibody responses to murine monoclonal antibody in mice. *J Pharm Sci*. 2015. doi:10.1002/jps.24413

28. Seefeldt MB, Rosendahl MS, Cleland JL, Hesterberg LK. Application of high hydrostatic pressure to dissociate aggregates and refold proteins. *Curr Pharm Biotechnol*. 2009;10(4):447-455.

29. Jiskoot W, Kijanka G, Randolph TW, et al. Mouse Models for Assessing Protein Immunogenicity: Lessons and Challenges. *J Pharm Sci*. 2016;105(5):1567-1575. doi:10.1016/j.xphs.2016.02.031

30. Ryff JC, Schellekens H. Immunogenicity of rDNA-derived pharmaceuticals. *Trends Pharmacol Sci*. 2002;23(6):254-256. doi:10.1016/S0165-6147(02)02024-2

31. Hochuli E. Interferon immunogenicity: Technical evaluation of interferon alpha 2a. *J Interf cytokine Res*. 1997;17:S15-S21.

32. Prummer O. Endogenous antibodies directed against interferon-alpha: paradigm of the immunogenicity of cytokines and its clinical impact. *Med Welt*. 1998;49(6):267-274.

33. Ring J, Stephan W, Brendel W. Anaphylactoid Reactions to Infusions of Plasma Protein and Human-Serum Albumin - Role of Aggregated Proteins and of Stabilizers Added during Production. *Clin Allergy*. 1979;9(1):89-97.

34. Moore W V., Leppert P. Role of Aggregated Human Growth-Hormone (Hgh) in Development of Antibodies to Hgh. *J Clin Endocrinol Metab*. 1980;51(4):691-697.

35. Ahmadi M, Bryson CJ, Cloake EA, et al. Small amounts of sub-visible aggregates enhance the immunogenic potential of monoclonal antibody therapeutics. *Pharm Res*. 2015;32(4):1383-1394. doi:10.1007/s11095-014-1541-x

36. Joubert MK, Hokom M, Eakin C, et al. Highly aggregated antibody therapeutics can enhance the in vitro innate and late-stage T-cell immune responses. *J Biol Chem*. 2012;287(30):25266-25279. doi:10.1074/jbc.M111.330902

37. Rombach-Riegraf V, Karle AC, Wolf B, et al. Aggregation of human recombinant monoclonal antibodies influences the capacity of dendritic cells to stimulate adaptive T-cell responses in vitro. *PLoS One*. 2014;9(1). doi:10.1371/journal.pone.0086322

38. Barandun S, Jeunet F, Kistler P, Isliker H. Intravenous Administration of Human Gamma-

Globulin. *Vox Sang*. 1962;7(2):157-174.

39. Ellis EF, Henney CS. Adverse reactions following administration of human gamma globulin. *J Allergy*. 1969;43(1):45-54. doi:10.1016/0021-8707(69)90019-7

40. Joubert MK, Luo Q, Nashed-Samuel Y, Wypych J, Narhi LO. Classification and characterization of therapeutic antibody aggregates. *J Biol Chem*. 2011;286(28):25118-25133. doi:10.1074/jbc.M110.160457

41. Bessa J, Boeckle S, Beck H, et al. The immunogenicity of antibody aggregates in a novel transgenic mouse model. *Pharm Res*. 2015;32(7):2344-2359. doi:10.1007/s11095-015-1627-0

42. Maddux NR, Daniels AL, Randolph TW. Microflow Imaging Analyses Reflect Mechanisms of Aggregate Formation: Comparing Protein Particle Data Sets Using the Kullback-Leibler Divergence. *J Pharm Sci*. 2017;106(5):1239-1248. doi:10.1016/j.xphs.2017.01.030

43. Kotarek J, Stuart C, De Paoli SH, et al. Subvisible Particle Content, Formulation, and Dose of an Erythropoietin Peptide Mimetic Product Are Associated with Severe Adverse Postmarketing Events. *J Pharm Sci*. 2016;105(3):1023-1027. doi:10.1016/S0022-3549(15)00180-X

44. FDA. Highlights Of Perscribing Information: Omontys. 2012.

45. Kalonia C, Kumru OS, Prajapati I, et al. Calculating the mass of subvisible protein particles with improved accuracy using microflow imaging data. *J Pharm Sci*. 2015;104(2):536-547. doi:10.1002/jps.24156

46. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. 1987;20(C):53-65. doi:10.1016/0377-0427(87)90125-7

47. Dempster A, Laird N. *Maximum Likelihood from Incomplete Data via the EM Algorithm*. Vol 39.; 1977. http://www.ams.org/leavingmsn?url=http://links.jstor.org/sici?sici=0035-9246(1977)39:1%3C1:MLFIDV%3E2.0.CO;2-Z&origin=MSN.

48. Zölls S, Gregoritza M, Tantipolphan R, et al. How Subvisible Particles Become Invisible — Relevance of the Refractive Index for Protein Particle Analysis. *Pharm Biotechnol*. 2013;102(5):1434-1446. doi:10.1002/jps

49. Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions. *IEEE Conf Comput Vis Pattern Recognit*. 2015. doi:10.1109/CVPR.2015.7298594

50. Bojarski M, Del Testa D, Dworakowski D, et al. End to End Learning for Self-Driving Cars. 2016:1-9. http://arxiv.org/abs/1604.07316.

51. Arindra A, Setio A, Ciompi F, et al. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. *IEEE Trans Med Imaging*. 2016;35(5):1160-1169. doi:10.1109/TMI.2016.2536809

52. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118. doi:10.1038/nature21056

53. Strehl R, Rombach-Riegraf V, Diez M, et al. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: A novel multiparametric image filter for sub-visible particles in microflow imaging analysis. *Pharm Res*. 2012;29(2):594-602. doi:10.1007/s11095-011-0590-7

54. Zölls S, Weinbuch D, Wiggenhorn M, et al. Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments. *AAPS J*. 2013;15(4):1200-1211. doi:10.1208/s12248-013-9522-2

55. Saggu M, Patel AR, Koulis T. A Random Forest Approach for Counting Silicone Oil Droplets and Protein Particles in Antibody Formulations Using Flow Microscopy. *Pharm Res*. 2017;34(2):479-491. doi:10.1007/s11095-016-2079-x

Chapter 4: Machine Learning & Statistical Analyses for Extracting and Characterizing "Fingerprints" of Antibody Aggregation at Container Interfaces from Flow Microscopy Images

Published as: Daniels AL, Calderon CP, Randolph TW. Machine Learning & Statistical Analyses for Extracting and Characterizing “Fingerprints” of Antibody Aggregation at Container Interfaces from Flow Microscopy Images. Biotechnol Bioeng. 2020.

4.1 Abstract

Therapeutic proteins are exposed to numerous stresses during their manufacture, shipping, storage and administration to patients, causing them to aggregate and form particles through a variety of different mechanisms. These varied mechanisms generate particle populations with characteristic morphologies, creating “fingerprints” that are reflected in images recorded using flow imaging microscopy. Particle population fingerprints in test samples can be extracted and compared against those of particles produced under baseline conditions using an algorithm that combines machine learning tools such as convolutional neural networks with statistical tools such as nonparametric density estimation and Rosenblatt transform-based goodness-of-fit hypothesis testing. This analysis provides a quantitative method with user-specified Type 1 error rates to determine whether the mechanisms that produce particles in test samples differ from particle formation mechanisms operative under baseline conditions. As a demonstration, this algorithm was used to compare particles within intravenous immunoglobulin (IVIg) formulations

that were exposed to freeze-thawing and shaking stresses within a variety of different containers. This analysis revealed that seemingly subtle differences in containers (e.g., glass vials from different manufacturers) generated distinguishable particle populations after the stresses were applied. This algorithm can be used to assess the impact of process and formulation changes on aggregation-related product instabilities.

4.2 Introduction

Aggregation is a major challenge in the manufacturing of therapeutic proteins[1–3]. Numerous stresses encountered during protein production cause aggregation. These different stresses (e.g., freeze-thawing[4–6], interactions at air-water and container-water interfaces[7–11], exposure to excipient degradation products such as those from polysorbates[12–14], pH extremes[15,16], and elevated temperatures) produce polydisperse distributions of aggregates[17]. As a result, aggregates may be observed in protein formulations following purification[18], filtration[19–21], pumping[22–24], freezing[5,25–27], vial filling[28], viral clearance steps and shipping[29]. The potential role of these aggregates in provoking unwanted immune responses[30–34] has generated interest in developing techniques to identify their root causes.

The root cause of protein aggregation is often elusive. However, the various stresses that promote protein aggregation each induce aggregation by somewhat different molecular mechanisms[35,36]. These distinct mechanisms lead to particle populations whose size and morphology distributions comprise particle “fingerprints” that reflect the root cause of their formation. Better techniques for characterizing these particle fingerprints would provide methods to rapidly determine the root causes of particle formation in a sample.

Flow imaging microscopy (FIM) is a commonly used technique for analyzing size distributions of protein aggregates[37–40] and other particles. FIM uses light microscopy combined with microfluidics to capture digital images of particles larger than one micron in size contained within a sample. The output from this instrument is a set of digital images of individual particles in a small liquid sample (usually about 10^3-10^5 images per 200 µL sample). The images contain a large amount of morphological information. However, in common practice most of the morphology information potentially available from FIM measurements is not utilized.

Convolutional neural networks (ConvNets) can be used to extract and analyze morphological information embedded in FIM images[41,42]. ConvNets are a family of neural networks capable of learning relevant features from a collection of images that are useful when performing tasks such as classification and dimension reduction[41,43–45]. ConvNets trained on FIM datasets can accurately classify protein aggregates produced by different stresses. In Calderon *et al.* 2018 and Gambe-*Gilbuena et al*. 2020, a set of single, well-defined stresses (e.g., freeze-thawing, heating) was applied to protein solutions, causing aggregates to form. ConvNets were then trained on FIM images of the resulting particles in order to train classifiers to recognize particle morphologies generated by one of these stresses. The resulting classifiers were then used to classify FIM images of particles from new samples that had been subjected the same set of stresses.

Although these previous approaches are useful for analyzing protein aggregates within formulations exposed to single stresses, protein aggregates encountered in practice are likely the result of a superposition of a variety of stresses, yielding more varied fingerprints. The potentially large number of different aggregate sources may mask subtle but relevant changes in particle populations due to minor changes in process conditions such as changes in container-closure systems. This issue is compounded by the inherent variability in particle

morphology even under tightly controlled conditions[42]. Thus, it can be difficult to determine if morphology differences within a particle population reflect different root causes of aggregation or merely sample-to-sample variability.

In the present study, we demonstrate a ConvNet algorithm that can be used to quantitatively determine if particle morphologies recorded in a small collection of FIM images are statistically different from those generated under a user-defined baseline condition. This analysis uses a combination of dimension reduction and hypothesis testing. Facial recognition strategies[46,47] such as triplet loss approaches[45] can reduce the dimensionality of FIM image data sets, compressing the information contained in color FIM images to two-dimensional feature vectors (i.e. the fingerprints). The extreme information compression enables the use of nonparametric techniques such as kernel density estimates of the probability density of these low-dimensional representations for particles made under a single baseline condition. Goodness-of-fit hypothesis test with user-tunable false positive rates can then be used to compare collections of particle images from other samples to this density.

One potential application of this approach is testing whether formulation design decisions (e.g. pH, excipient concentrations, container-closure types) affect protein aggregate populations. In this work, we focus on the impact of container-closure systems on protein aggregate morphology. The geometry and chemistry of the container can affect protein aggregation[48]. Container-induced particles may come directly from the container (e.g. glass flakes from delamination in glass vials[49]) as well as from protein aggregates triggered by the container itself[50,51]. Aggregation may depend not only on the type of container[52–54] but also may vary between different lots of the same container from a given manufacturer.

4.3 Materials and Methods

4.3.1 Materials

Intravenous immunoglobulin (IVIg; GAMMAGARD LIQUID) was obtained from Takeda International (Lexington, MA). Phosphate buffered saline (PBS) containing 144 mg/L potassium phosphate monobasic, 795 mg/L potassium phosphate dibasic, and 9000 mg/L sodium chloride at pH 7.4 was obtained from Gibco (Waltham, MA). Polypropylene, 2 mL microcentrifuge tubes ("Plastic") were from Fisher Scientific (Waltham, MA). FIOLAX Clear 3 mL Type 1 borosilicate glass vials ("Glass 1") were obtained from Schott (Elmsford, NY). A second 3 mL Type 1 borosilicate glass vial ("Glass 2") was obtained from Duran Wheaton Kimble (Mainz, Germany). Micro-90 was obtained from International Products Corp. (Burlington, NJ). Polystyrene 20-µm calibration beads were from Thermo Scientific (Waltham, MA).

4.3.2 Generation of Protein Aggregates

IVIg aggregates were made using combinations of two aggregation-inducing stresses in three container types. Five experimental replicates were made per combination of container and stress. In each replicate, two containers were cleaned by filling the container with ultrapure water generated using a PURELAB flex 1 water deionization system from ELGA Labwater (Wycombe, UK), shaking the filled container, then emptying the container and allowing the container to air dry for one hour. IVIg stock solution was made by centrifuging the as-received drug product containing 100 mg/mL IVIg at 15,000 g for 20 minutes at 4°C. in. The supernatant was then diluted to 0.5 mg/mL using filtered PBS, and 1.5 mL of this solution was filled into each

container. Samples were then exposed to either freeze-thawing or shaking stresses as described below.

Freeze-Thaw Stress: Samples stressed by freeze-thawing underwent four freeze-thaw cycles. During each cycle, the samples were suspended in a fixed orientation in liquid nitrogen for 4 minutes and then suspended in a 30°C water bath for 10 minutes. FIM analysis was performed immediately after the final freeze-thaw cycle was completed.

Shaking stress: Samples were taped in a horizontal orientation onto an orbital plate shaker and shaken at 800 RPM for 4 hours. FIM analysis was performed immediately after the shaking was completed.

4.3.3 Flow Imaging Microscopy

FIM images were recorded with a FlowCam® VS instrument (Fluid Imaging Technologies, Inc., Scarborough, ME) with a 10x objective, a field-of-view flow cell with a depth of 80 µm and width of 700 µm, and color imaging. The instrument was focused using the built in autofocus protocol for optimal image quality using 20-µm calibration beads. 1% Micro-90 solution followed by filtered ultrapure water were flushed through the instrument before and between measurements. The flash duration of the instrument was adjusted between replicates to achieve a constant background intensity of 150. Three 0.2 mL aliquots were analyzed from each replicate vial. Images were collected at a flow rate of 0.05 mL/min using 15 light and 17 dark pixel thresholds for particle segmentation.

4.3.4 Image postprocessing

FIM images of particles were imported into Python 2.7. Before further analysis, the size of each image was adjusted to 24 x 24 pixels. Smaller images were padded with pixels sampled from a normal distribution with the same mean and variance as the border of the image and smoothed using Gaussian smoothing. For larger images a centered 24 x 24 crop of the image was used. Three experimental replicates for each combination of container and stress were used to train the algorithm, while the remaining two independent replicates were retained for use in subsequent testing. 14,000 images were randomly selected from each of the three training replicates to be used as training data for the algorithm described in the next section. The remaining two replicates for each condition were not shown to the algorithm at all during training. 2,000 images from each replicate, including those not included in algorithm training, were set aside during algorithm training and used to test the performance of the trained algorithm.

4.3.5 Algorithm Overview

An algorithm was developed to determine if FIM images from a test sample were statistically consistent with those in a baseline sample. Hereafter, we refer to these FIM images as “particles” since each FIM image is recorded on a single particle. Figure 9 shows the process of training the algorithm to identify particles in a baseline sample. First, a convolutional neural network (ConvNet) is trained on the collection of FIM images (Figure 9, first row, first column) to compress information within these images into a low-dimensional (2D here) point cloud of embeddings (Figure 9, second row, first column). A nonparametric kernel density estimate is then constructed from this low-dimensional point cloud to estimate the probability density of

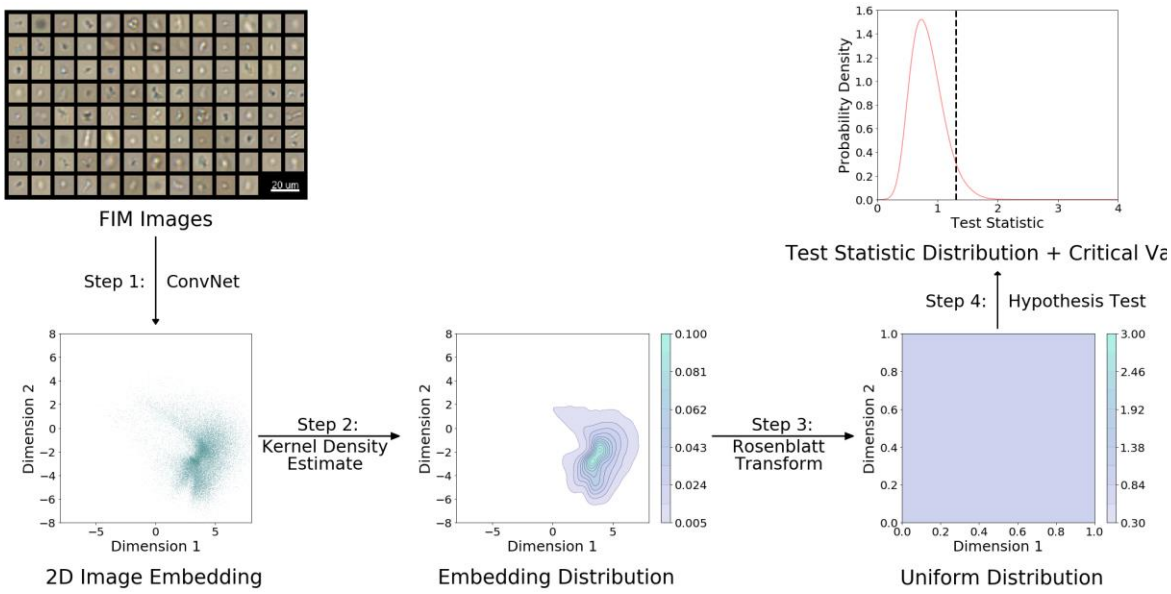

Figure 9: Particle comparison algorithm training flow chart
Flow chart showing how the algorithm is trained to detect particle populations similar to those made under some baseline condition. FIM images of particles made under the baseline condition (first figure) are used to train a ConvNet capable of compressing the image into a two-dimensional embedding (second figure). The probability density of these embeddings is then estimated using a kernel density estimate (third figure). A Rosenblatt transform defined using this distribution can then be used to map embeddings from the baseline sample onto a uniform distribution (fourth figure). The transformed embeddings can then be used to set up goodness-of-fit hypothesis tests by estimating the distribution of a test statistic for the baseline sample (fifth figure, curve) and using the distribution to determine an appropriate critical value for the test (fifth figure, dashed line).

embeddings in the baseline sample (Figure 9, second row, second column). The estimated probability density is subsequently used to define a Rosenblatt transform which maps an embedding to a new random vector having the same dimensions of the embedding point [55] (Figure 9, second row, third column). Goodness-of-fit hypothesis tests can be applied in conjunction with this Rosenblatt transform to determine if sets of FIM image embedding points are consistent with the estimated baseline density. The hypothesis test exploits the following mathematical fact: if a collection of embedding points are distributed according to the probability density associated with the baseline sample, the Rosenblatt transform yields multivariate random vectors whose components are independent and identically distributed with each component being a uniformly distributed random variable between 0 and 1. Goodness-of-fit

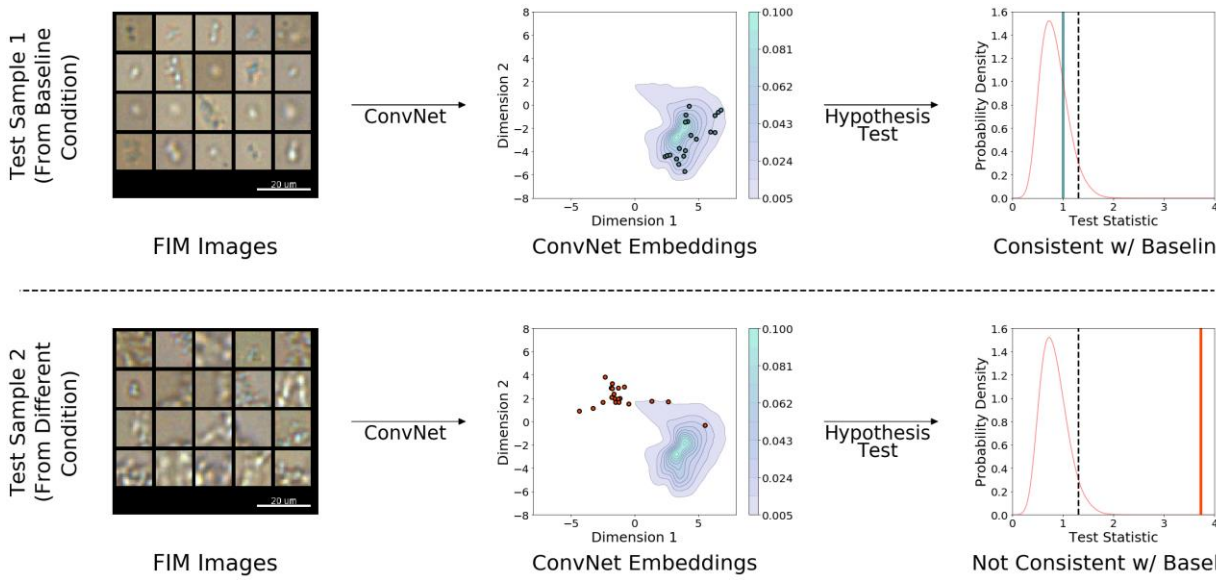

Figure 10: Particle comparison algorithm testing flow chart
Flow chart showing the application of the algorithm to test samples that either resemble the baseline sample (top row) or do not resemble the baseline sample (bottom row). To perform the analysis small sets of images are selected from each sample (first column) and analyzed with the ConvNet to obtain 2D embeddings for the images (second column, points). These embeddings are then compared against the distribution of embeddings for the baseline sample (second column, contour) using a combination of Rosenblatt transforms and hypothesis testing (third column). If the test statistic for the test sample (third column, solid line) is less than critical value for the baseline sample (third column, dashed line), the algorithm identifies that the particles in the test sample are consistent with the baseline sample.

hypothesis testing can formally check sets of transformed embeddings for this property. Critical values for this hypothesis test are set by repeatedly subsampling Rosenblatt-transformed embedding points from the baseline sample, calculating test statistics for each subsample, selecting a value based on the resulting test statistic distribution to obtain a user specified Type I error (i.e. false positive) rate (Figure 9, first row, third column)

Once trained, the algorithm can be used to quantify how statistically similar particle populations in test samples are to that in the baseline sample. Figure 10 shows the application of the trained algorithm to test samples containing either similar or different particle populations. To analyze a

test sample, a small number of FIM images (e.g. 5-200) are subsampled from the test sample, converted to 2D embeddings with the trained ConvNet, and transformed using the Rosenblatt

Transform defined by the baseline density. Goodness-of-fit hypothesis tests using the critical values from the baseline sample are then used to test if the transformed embeddings are consistent with a uniform distribution. Applying this algorithm to particles that resemble those in the baseline sample (Figure 10, top row) results in embeddings that are both visually and statistically consistent with those in the baseline sample. Conversely, particles that do not resemble the baseline sample (Figure 10, bottom row) yield embeddings less consistent with the baseline sample and are thus identified through goodness-of-fit hypothesis testing as a different particle population from the baseline sample.

4.3.6 Convolutional Neural Networks (ConvNets)

Convolutional neural networks (ConvNet) are used in this analysis to extract and compress information in FIM images into a set of image features. While previously these image features were used as the input to a classifier that predicted the stress to which a sample had been exposed[41], in the current analysis nonparametric techniques were used to estimate the distribution of these features. To apply these techniques, the ConvNet needed to be trained to learn extremely low-dimensional (i.e. 2-3 image features) representations of FIM images in order to avoid the exponential decrease in accuracy of these techniques with each additional dimension in the data[56].

The ConvNet used in this analysis was trained using a triplet loss algorithm, an approach that was developed for facial recognition to learn highly-compressed image representations[45]. In this

algorithm, a ConvNet is trained to learn a low-dimensional representation or embedding of images that acts to cluster together images from similar sources (e.g. faces of the same person, or protein aggregates made by the same stress and in the same container). During training, triplets (i.e. sets of three images) are assembled from the training data consisting of an image of one particle type (the anchor image), another image of the same particle type (a positive image), and a third image of a different particle type (the negative image). These triplets are fed through the neural network to calculate embeddings for each of the three images. The network's parameters are then adjusted to minimize a modified triplet loss function[57]:

$$l = \log\left(\exp\left(-(d_{ap} - d_{an} + \alpha)\right) + 1\right) \quad (1)$$

Where l is the triplet loss, d_{ap} is the Euclidian distance between the representations of the anchor and positive images returned by the ConvNet, d_{an} the distance between the representations of the anchor and negative image, and α is the margin, a small number used to scale the distances between dissimilar particle types in the embedding. This loss function is minimized when particles from a common source are close to each other in the embedding space and far apart from particles from other sources. In addition to allowing nonparametric density estimation techniques to be used, the resulting ConvNet can also be used to effectively analyze FIM image types not shown to the network during training. The algorithm is trained on a modest number (e.g. >10,000) of particles made under a small set conditions, but once trained, the network can then be used to analyze conditions not shown to the network during training using a smaller number of particles.

A ConvNet was trained to compress the particle information in 24 x 24 x 3 preprocessed FIM images into a 2D representation of the image. The network structure used in this study is shown

Table 3: ConvNet structure used in glass comparison study

Layer	Layer Type	# of Features	Kernel Size	Activation	Input Shape	Output Shape
1	Convolutional	32	3x3	ReLU	24x24x3	22x22x32
2	Convolutional	32	3x3	ReLU	22x22x32	20x20x32
3	Dropout (10% rate)	-	-	-	20x20x32	20x20x32
4	Convolutional	32	3x3	ReLU	20x20x32	18x18x32
5	Convolutional	64	3x3	ReLU	18x18x32	16x16x64
6	Max Pooling (2x2)	-	-	-	16x16x64	8x8x64
7	Dropout (10% rate)	-	-	-	8x8x64	8x8x64
8	Convolutional	64	3x3	ReLU	8x8x64	6x6x64
9	Convolutional	64	3x3	ReLU	6x6x64	4x4x64
10	Flatten	-	-	-	4x4x64	1,024
11	Dropout (10% rate)	-	-	-	1,024	1,024
12	Dense	64	-	Softplus	1,024	64
13	Dense	2	-	None	64	2

in Table 3. The Visual Geometry Group (VGG)-inspired network[58] uses convolutional layers with rectified linear unit (ReLU) activations in conjunction with max pooling and dropout layers. Unlike the VGG structure, the first dense layer of the network used softplus activations[59] as the sparsity introduced by ReLU activations was found to cause non-smooth features in embedding point clouds which subsequently complicated obtaining accurate kernel density estimates of the probability density of the points.

The ConvNet was trained on FIM images of particles produced in Plastic and Glass 1 vials after applying either freeze-thaw or shaking stresses. Particles generated within Glass 2 vials were not used to train the ConvNet, but instead were used to test the network's generalization to unseen particle types. The network was trained with a margin of 0.5 using minibatches of 64 triplets using an Adam optimizer[60] with a 0.001 learning rate. Triplet minibatches were generated by assembling minibatches of 64 anchor images from the training images and calculating image embeddings for each training image at the start of each epoch. Positive and

negative images for each anchor image were then randomly selected from all training images until a triplet was found that met semi-hard triplet mining criteria[45] based on the most recently-calculated embeddings. This approach filters out triplets that have low and high values of the loss function which can prevent the network from learning effective image representations. Image augmentation including flipping images horizontally and vertically 50% of the time as well as adding a random RGB value sampled from a normal distribution with zero mean and 0.01 standard deviation to all pixels in the image was performed. The current value of the triplet loss function as well as the variance in embeddings from each condition was monitored during training at the end of each epoch. The network was trained for 100 epochs and the network parameters that minimized the triplet loss was used in subsequent steps of the analysis.

4.3.7 Kernel Density Estimation

Kernel density estimation is a nonparametric technique for estimating the probability density function (PDF) of a dataset using data sampled from this distribution[56]. This technique was used to estimate the distribution of the low-dimensional FIM image embeddings for the baseline sample directly from the embeddings. Embedding sets from test samples were then compared against this distribution to decide if the particles in the test sample were consistent with those in the baseline sample.

Kernel density estimates of the distribution of embeddings for the baseline sample were constructed using a product kernel and using normal distributions as the kernel in each dimension. This kernel function was chosen so that the estimated PDF has an infinite support, which was helpful in obtaining meaningful evaluations of the Rosenblatt transform on particles

that embedded far away from the mode of the PDF. The bandwidth of the kernel in each dimension was calculated using a normal reference rule[56]:

$$h_i = \left(\frac{4}{d+2}\right)^{1/d+4} \sigma_i n^{-1/d+4} \quad (2)$$

where h_i is the bandwidth in dimension i, d is the number of dimensions of the embeddings (2 in this study), n is the number of datapoints used to construct the density estimate and σ_i is the standard deviation of the embeddings in dimension i.

4.3.8 Rosenblatt Transform

The Rosenblatt Transform is a statistical normalization that maps a d-dimensional random variable onto a new d-dimensional random variable using the PDF believed to govern the original variable[55]. If the PDF does govern the initial random variable, the resulting random variable will be governed by a d-dimensional uniform distribution in which each dimension is independent and identically distributed. Goodness-of-fit hypothesis testing can then be used to check if the transformed variable is consistent with the uniform distribution. Performing this testing on the transformed variable allows us to use simple goodness-of-fit hypothesis test statistics with known distributions for a uniform null PDF—regardless of the initial null PDF that would be used if the testing was performed on the untransformed embeddings. The Rosenblatt transform also ensures that the two dimensions of the embeddings are independent under the null hypothesis. This feature of the transform simplifies setting critical values for the hypothesis test as the critical value does not need to be adjusted to account for dependencies in the embedding dimensions under the null hypothesis. Embedding sets from test samples were

transformed based on the embedding PDF for the baseline sample prior to performing goodness-of-fit hypothesis testing.

4.3.9 Goodness-of-Fit Hypothesis Testing

After applying the Rosenblatt Transform defined by the density estimate for the baseline sample to image embeddings from the test sample, goodness-of-fit hypothesis tests were used to test the null hypothesis that the transformed embeddings are consistent with a uniform distribution. Rejection of this null hypothesis indicated that the particles in the test sample were not consistent with those in the baseline sample and thus potentially formed under a different set of conditions.

All goodness-of-fit hypothesis testing was performed using a Kolmogorov-Smirnov (KS) test,[61] a 1D goodness-of-fit hypothesis test that compares the cumulative distribution function (CDF) of the embeddings to the CDF of a 1D uniform distribution. Two 1D KS tests were performed on each dimension of the transformed embeddings to test the null hypothesis as the dimensions of the transformed embeddings are independent under the null hypothesis. The null hypothesis was rejected if either dimension was not consistent with a uniform distribution. The desired overall Type I error (i.e. false positive) rate of the test can be used to set error rates for the two individual tests obtained using the Bonferoni correction. The overall null hypothesis was rejected if either of the two tests rejected the null hypothesis. A 5% overall Type I error rate was used in this analysis.

Test statistics for the hypothesis test were calculated using sets of a small number (e.g. 5-200) of particles randomly selected from the test sample. Since the statistical power of these

hypothesis tests scales with the number of data points used in the analysis, restricting the number of particles that are analyzed at once helped control the sensitivity of the analysis. In this study sets of either 20 or 200 particles were used to compare test samples to the baseline sample.

Monte Carlo simulations were used to select appropriate critical values (aiming at obtaining a user-specified type I error rate) of the test statistic to account for the bias introduced by both the nonparametric density estimate and subsampling scheme. The test statistic distribution for sets of 20 particles was estimated by randomly subsampling 10,000 sets of 20 training particles from the baseline condition and evaluating the test statistics for each subsample. These distributions were then used to select critical values at the appropriate significance level for each test. This process was repeated using sets of 200 particles, resulting in a second test statistic distribution and critical value for these larger particle sets.

Test samples were compared against baseline samples by repeatedly subsampling sets of particles from the test sample and using the algorithm to identify the fraction of these subsamples that were consistent with the baseline sample. 2,500 sets of either 20 or 200 particles were subsampled from the test sample. After computing the Rosenblatt transformed embeddings for each subsample, the hypothesis test was used to determine if each subsample was consistent with the baseline sample. The similarity between the test sample and the baseline sample using a given number of particles was recorded as the fraction of the 2,500 subsamples that did not contain particles consistent with the baseline sample.

4.3.10 Particle Morphology Comparison

The algorithm described above was used to compare samples of aggregated IVIg formed under different stress conditions and in different containers. A ConvNet was trained on FIM images of particles made in Plastic and Glass 1 containers with a triplet loss approach. The remaining steps of the algorithm were then trained to identify sets of either 20 or 200 particles that resemble sets of the corresponding number of particles made in one container after exposure to one stress. These later steps were separately trained twelve times to cover the six possible baseline classes (particles made by one stress in one container) and the two particle set sizes (20 or 200) that were used during testing.

The trained algorithms were used to investigate the impact of different stresses and different containers on particle morphology. This comparison was performed by comparing small sets of test particles from each sample to all the training particles from a single container and stress. To investigate the impact of stresses on particle populations, the algorithm was used to compare particles generated by freeze-thaw stress (the baseline stress class) in each of the three containers to particles made in the same container type after exposure to shaking and freeze-thaw stresses. Similarly, the effect of container on particle populations was investigated by comparing particles made in Glass 1 containers after exposure to each stress to samples of those made in each of the three containers after exposure to the same stress.

4.3.11 Surface characterization

The two glass vial types used in this analysis were characterized using contact angles and surface profilometry. Each of these measurements was performed by cutting off the bottom of the vial and cleaning the inner surface with ethanol, water, and nitrogen before measurements.

To assess the hydrophobicity of the glass containers, contact angles were measured for each vial using a ramé-hart Model 210 goniometer / tensometer with DROPimage Pro software (Succasunna, NJ). This instrument was used to measure static, advancing, and receding contact angles on each of the three surfaces. These measurements were performed in triplicate. Between measurements, the surfaces were cleaned with ethanol, water and nitrogen gas.

Surface profilometry was performed with a Dektak 3030 Profilometer (Billerica, MA) to measure the roughness of the two vials. Surface profiles were measured along a flat 1 mm length along the inner surface of the vial. These profiles were fitted to a second-degree polynomial which was then subtracted from the raw data in order to account for the macroscopic curvature of these surfaces. The flattened surface profiles were then used to calculate the arithmetic average roughness R_a of each container which is calculated using:

$$R_a = \frac{1}{n}\sum_{i=1}^{n}|h_i| \tag{3}$$

Where n is the total number of locations along the 1 mm length that the height was measured, i indexes the different height measurements and h_i is the height measured at point i.

4.4 Results

4.4.1 Flow Imaging Microscopy

Figure 11 shows collections of randomly selected FIM images obtained from each of the six conditions compared in this analysis. These images reveal obvious differences between particles generated by freeze-thawing and shaking stresses; particles observed after shaking (figure 11, a-c) are typically large and exhibit complex morphologies while particles imaged after freeze-thaw cycling (figure 11, d-f) are much smaller with simple morphologies. Conversely, the effect of different containers on particle morphologies generated by these stresses is not visually obvious from the images.

4.4.2 Convolutional Neural Network:

Figure 12 shows contour plots of the distribution of embeddings returned by the trained ConvNet for particles made by freeze-thawing and shaking stresses. The contours for the different stresses are visually separated within this embedding space, indicating that the network can distinguish between particles generated by shaking and freeze-thaw stresses. Figure 12 also shows sample particles that are mapped to different locations in the embedding space. In this embedding scheme, small particles with simple but common structures are mapped near the mode of the freeze-thaw distribution whereas large, complex heterogeneous particles are mapped near the mode of the shaking distribution.

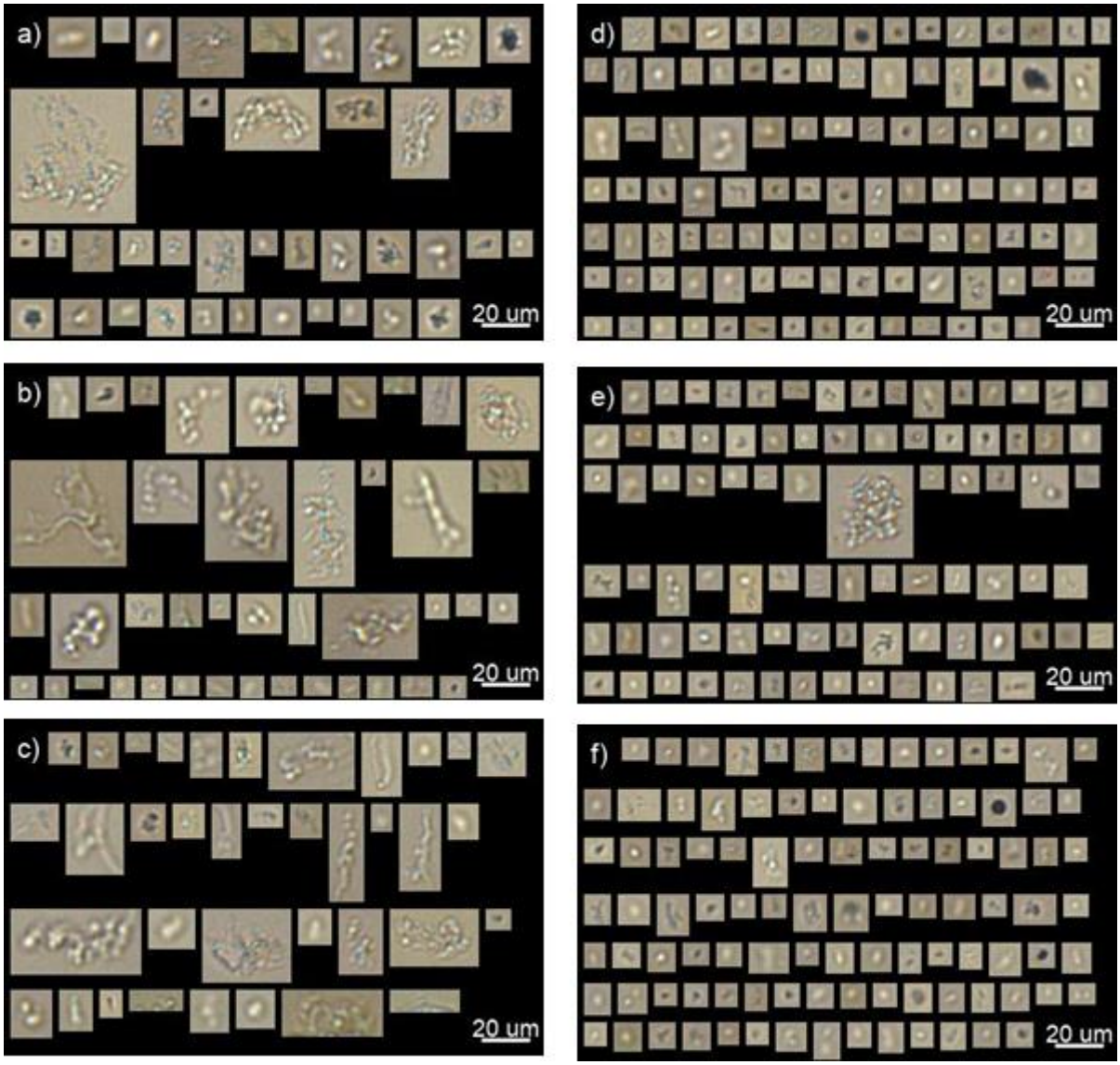

Figure 11: Sample container and stress FIM images
Sample collages of IVIg aggregates generated by (a-c) shaking and (d-f) freeze-thaw stress in (a, d) Plastic, (b, e) Glass 1 or (c, f) Glass 2 vials

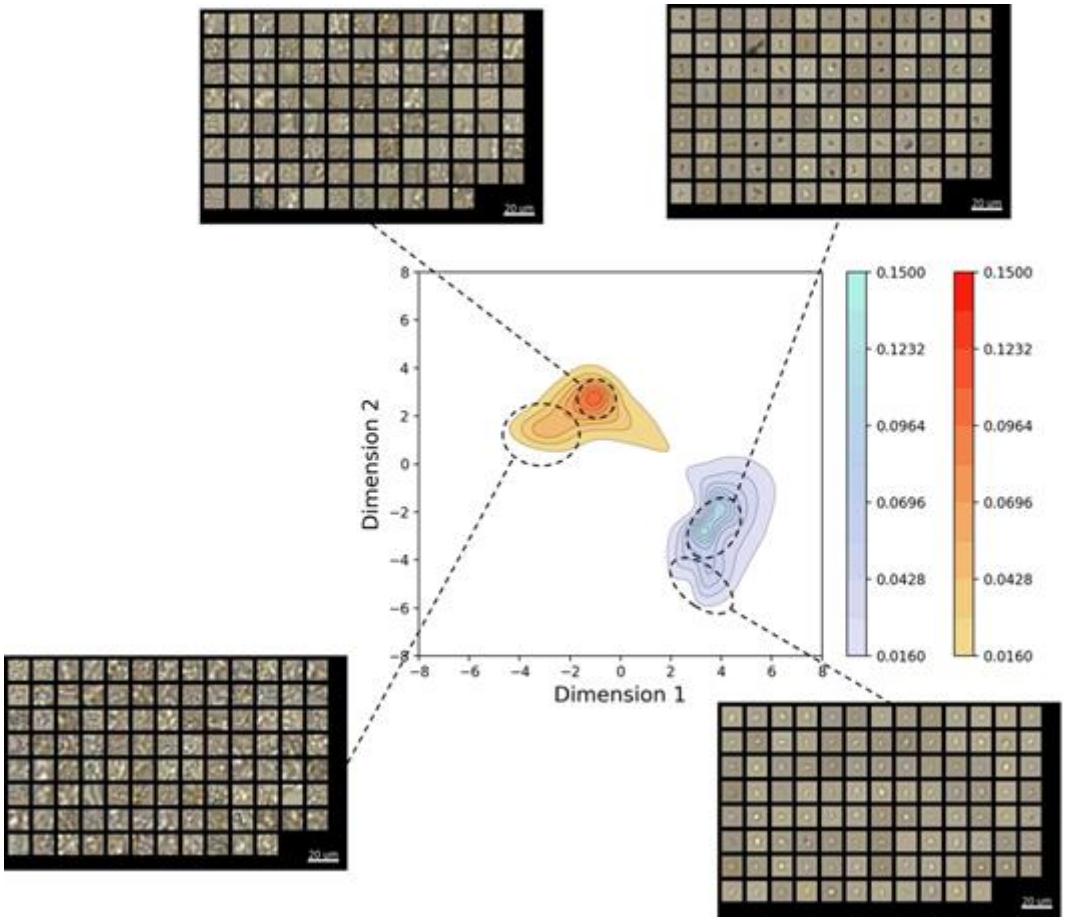

Figure 12: Stress and container embedding contours and collages
Contour plot of the FIM image embeddings of particles used to train the ConvNet. The colors of the plot indicate the density (i.e. PDF value) of image embeddings from a specific stress in that region of the embedding space. The darker orange-to-red contour corresponds to particles formed using shaking stress while the lighter blue-to-purple contour corresponds to particles formed using freeze-thaw stress. Regions of a single color have PDF values between the values indicated on the color bar. The values of the color transitions were manually selected to aid in viewing these PDFs. Also shown on each figure are typical FIM images that embed within specific regions of the contour plot.

Figure 13 shows contour plots of the estimated PDF of embeddings returned by the trained ConvNet for all particles not included in the network--including those from the four samples per condition that were not used to train the network. Figure 13(a) shows the embeddings for particles formed in Plastic containers, figure 13(b) shows those formed in Glass 1 containers, and figure 13(c) shows those formed in Glass 2 containers. These contour plots indicate an observable difference in the particle morphologies produced in the three containers as a result

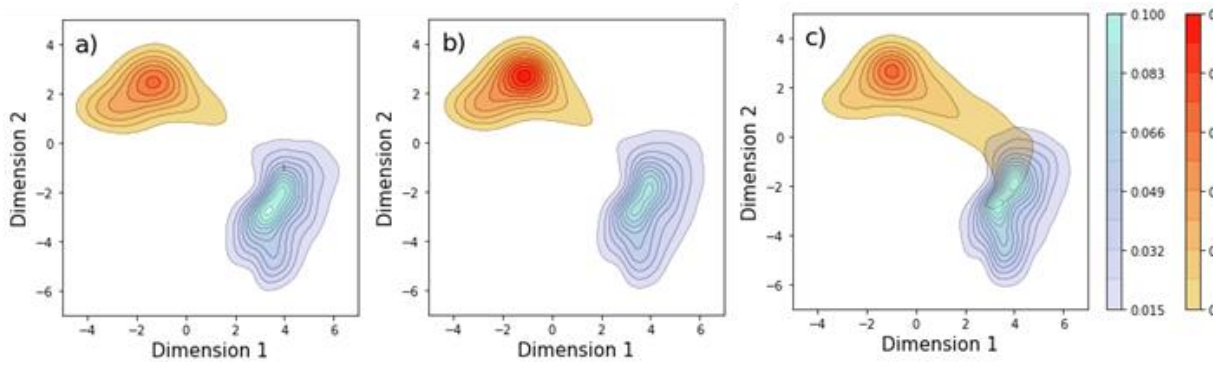

Figure 13: Stress and container effects on FIM image embeddings
Contour plots of the FIM image embeddings of particles not shown to train the algorithm made in (a) Plastic microcentrifuge tubes (b) Glass 1 vials and (c) Glass 2 vials. The darker orange-to-red contours in each plot correspond to particles made using shaking stress while the lighter blue-to-purple contours in each correspond to particles made using freeze-thaw stress. Colors in this figure are interpreted as described for figure 12

of shaking stresses. Compared to particles produced by shaking Glass 2 or Plastic containers, particles produced by shaking samples in Glass 1 containers (figure 13(b)) have a much tighter density in the embedding space than either of the other samples. Particles produced by freeze-thawing stress appear to be influenced to a lesser extent by the container in which they were formed; particles produced freeze-thaw cycling in Glass 1 containers exhibit a slightly more diffuse distribution than those produced in the other container types.

4.4.3 Particle comparisons

The remaining steps of the algorithm were used to compare the particle populations produced within different containers when exposed to different stresses. This comparison was done by choosing one of the samples being compared to be the baseline sample, subsampling small sets of particles from the other (test) sample, and testing the null hypothesis that each subsample contained particles that were consistent with those in the baseline sample. Table 4

Table 4: Stress comparison rejection rates

Probability that a set of 20 images of particles produced by freeze-thaw or by shaking stress in a container will be distinguishable from a baseline population of particles made by freeze-thaw stress in the same container

Baseline Sample: Particles Made by Freeze-Thaw Stress in Denoted Container	Probability of Rejecting Test Particle Sets (20 Particles)	
	Aggregates Made by Freeze Thaw Stress	Aggregates Made by Shaking Stress
Glass 1	5.0%	100.0%
Glass 2	5.9%	100.0%
Plastic	5.0%	100.0%

shows the rejection frequencies when comparing sets of 20 particles from each of the three containers to those made by applying freeze-thaw stress to the same container. As was expected from both the raw flow imaging data shown in figure 11 and the embeddings in figure 13, the algorithm can easily identify morphology differences between particle populations that had been exposed to these two stresses using only a small number of particle images. Additionally, the algorithm only mis-identified unseen test particles made under baseline conditions as being different from the baseline population around 5% of the time--the type I error rate that the test was designed to give.

Table 5 shows the rejection frequencies when comparing sets of either 20 or 200 test particles made by each stress to those produced in Glass 1 containers when exposed to the same stress. Interestingly, the ability of the algorithm to distinguish between particles produced in each of the three containers depended on the applied stress. Sets of 20 particles produced by exposing IVIG solutions in Plastic or Glass 2 containers to freeze-thawing stress were only able to be distinguished from those produced in Glass 1 containers at approximately the same rate as the Type I error rate. In contrast, sets of 20 particles produced by shaking stress in each container were distinguishable from those made in Glass 1 containers approximately 40% of the

Table 5: Container comparison rejection rates
Probability that a set 20 or 200 random particles formed in Glass 1, Glass 2, or Plastic containers by freeze-thaw or shaking stress will be distinguishable from a baseline population of particles made in the Glass 1 containers by the respective stress.

Baseline Sample: Particles Made in Glass 1 Containers by Denoted Stress	Number of Particles	Probability of Rejecting Test Particle Sets		
		Particles Formed in Glass 1 Containers	Particles Formed in Glass 2 Containers	Particles Formed in Plastic Containers
Freeze-Thaw	20	5.0%	6.8%	5.1%
	200	6.6%	40.6%	21.5%
Shaking	20	5.1%	36.0%	13.4%
	200	4.0%	100.0%	70.6%

time—eight times the type I error rate of the test. Increasing the size of the particle sets to 200 increased the fraction of shaking particle sets from Glass 2 and Plastic containers that were distinguished from those produced in Glass 1 containers with only a small increase in the false positive rate when the test was applied to held-out baseline samples. In addition, the larger particle sets allowed the algorithm to distinguish between particles made by freeze-thawing stress in Glass 1 and Glass 2 ~ 40% of the time and those produced in Glass 1 and Plastic 20% of the time.

4.4.4 Surface characterization

Table 6 shows the surface characterization results for the two types of glass vials. The contact angle measurements suggest that the two glasses have similar hydrophobicities, with Glass 1 being slightly more hydrophilic than Glass 2. Both surfaces were also found to have similar roughnesses.

Table 6: Contact angle and surface roughness measurements for Glass containers

Container	Static Contact Angle	Advancing Contact Angle	Receding Contact Angle	Contact Angle Hysteresis	Ra (nm)
Glass 1	16	23	<5	<18	28.
Glass 2	27	31	<5	<26	12

4.5 Discussion

This study presents and demonstrates a novel algorithm designed to compare FIM images of protein aggregates and other particles obtained from one sample to those obtained in some baseline sample. This approach is a departure from previous techniques used to predict to which of a small set of conditions a sample was exposed[41,42]. The primary advantage of this new approach is its ability to determine, *using only a small number of FIM images***,** if a new sample exhibits significantly different particle populations than those found under baseline conditions. The combination of traditional statistical tools with powerful machine learning algorithms can be used to determine if two samples exhibit a morphology difference that cannot be explained by sample-to-sample variance in particle morphology under a single root cause. This approach is effective at identifying (statistically) significant differences in particle morphology occurring due to different root causes such as manufacturing changes or process upsets that could warrant further investigation.

The use of statistical tools in this algorithm also give users control over the sensitivity of the analysis to changes in particle morphology. Decreasing the Type I error rate or increasing the number of particles used in the hypothesis test increases the sensitivity of the test so that smaller deviations in particle morphology from the baseline condition are identified as significant. This feature allows the sensitivity of the algorithm to be tuned for a specific application. For instance, the sensitivity of the algorithm can be increased for formulation

development in order to better detect subtle changes in particle morphology between possible formulations. In contrast, in process monitoring applications the sensitivity of the algorithm can be decreased to minimize the chances that a false positive difference between a baseline "normal" reference batch and a new batch of product triggers unnecessary process shutdowns.

The techniques used to learn low-dimensional FIM image representations, calculate density estimates, and perform goodness-of-fit hypothesis testing were chosen to demonstrate the algorithm using relatively simple techniques. While the techniques used here were effective in this analysis, in practice other techniques for these analyses could be considered in order to further improve the performance of the algorithm. For example, different goodness-of-fit hypothesis tests[62–64] may provide better statistical power against deviations in particle morphology than the test used here.

The performance of this algorithm was demonstrated on particles made by subjecting IVIg solutions in three types of containers to two different stresses. Freeze-thawing and shaking stresses produced particle populations that would be easy to distinguish by visual inspection of the FIM images (Figure 11). These stresses produced visually-resolved embeddings in the learned embedding space (Figures 12-13) which can then be easily distinguished using hypothesis testing (Table 4). It should be noted that this approach can still be used to classify samples by the stresses that they were exposed to as was done in previous papers[41,42]. If FIM images of different suspected stresses are available, this algorithm can be used to check if the particles in a sample are consistent with those produced by one of these stresses.

The algorithm was also able to identify the impact of container surfaces on particle populations. The results shown in Table 5 suggest that the particle populations produced by freeze-thaw and shaking stress are influenced by the container, but that the effect is more obvious when shaking

stress is used to create particles. The larger impact of the container on the particles produced by shaking stress agrees with the mechanistic understanding of these stresses. Agitation-induced aggregation likely occurs at interfaces including the air-water and container-water interfaces[51,54]. In contrast, during freeze thawing aggregation due to adsorption to ice-water interfaces and cryoconcentration effects[65] may occur at locations removed from container interfaces. Thus, the container interfaces might be expected to impact particle populations more when shaking stresses rather than freeze-thawing stresses are used to cause aggregation.

While the effect of container type was more subtle for particles made via freeze-thaw stress, the distribution of FIM image embeddings showed increased density near the mode of the distribution for Glass 2 vials as compared to that for Glass 1 vials (Figure 13). This difference was statistically discernible when sets of 200 particle images were analyzed. This result is somewhat surprising given the expected limited role of the container-water interface on aggregation induced by freeze-thawing. One possible explanation for the different particle fingerprints observed following freeze thaw cycling in the two types of glass vials is differences in heat transfer through the vial walls. Differences in the thickness or geometry of the glass between the two vial types could cause a difference in the heat transfer rate through the container walls. Higher heat transfer rates would accelerate the growth of ice crystals from the walls of the container that occurs during liquid nitrogen-induced freezing[66]. This faster growth results in a larger amount of ice interfacial area[67] and increased protein inclusion within growing ice crystals[6,68] which can induce protein unfolding and aggregation[69,70].

The algorithm was not explicitly trained to detect the observed differences between particle populations produced in Glass 1 and 2 containers as particles generated in Glass 2 were not used to train the ConvNet embedding step. The ability to compare unseen particle types against those in a user-defined baseline allows new samples to be analyzed using a fraction of the FIM

images (20-200) that would be required to retrain a ConvNet on a new particle type (>10,000). The required number of FIM images can be recorded rapidly using small volumes of sample.

The algorithm revealed that different types of particles can form in a single protein formulation when stressed in different types of containers, even when the containers are as similar as the two borosilicate glass container types tested here. The container-dependent formation of different particles would have been difficult to predict using simple surface characterization techniques, since the glasses have similar roughness and hydrophobicity (Table 6).

The analysis presented here can be used to compare the effect of changes in container types (e.g., new lots of glass vials) on protein stability using an approach that incorporates standard accelerated stability protocols. Prior to any change, a baseline set of FIM images should be obtained after subjecting the protein formulation to accelerated stability conditions (e.g. agitation, freeze-thawing), capturing images of the resulting particles using FIM and training the algorithm to recognize the imaged particles The accelerated stability protocol can then be repeated on a small number of containers in the new lot, and then the trained algorithm can be used to analyze whether new types of particles are associated with the new container lot. If the new container lot is found to produce statistically different particle populations, the lot may require additional characterization prior to use with the drug product

4.6 Acknowledgements

Funding for this work was provided by AstraZeneca and Department of Education GAANN Award P200A150211. CC was funded by NIH Award R41GM130513. The authors (AD, CC, and TR) are inventors on intellectual property related to this manuscript that is owned by the

Regents of the University of Colorado and Ursa Analytics. This work was sponsored in part through an STTR grant to Ursa Analytics, a company in which CC has a controlling financial interest.

4.7 References

1. Randolph TW, Carpenter JF. Engineering Challenges of Protein Formulations. *AIChE J*. 2007;53(8):215-228.

2. Roberts CJ. Protein aggregation and its impact on product quality. *Curr Opin Biotechnol*. 2014;30:211-217. doi:10.1016/j.copbio.2014.08.001

3. Wang W. Instability, Stabilization, and Formulation of Liquid Protein Pharmaceuticals. Vol 185.; 1999.

4. Arsiccio A, Pisano R. Stability of Proteins in Carbohydrates and Other Additives during Freezing: The Human Growth Hormone as a Case Study. *J Phys Chem B*. 2017;121:8652-8660.

5. Barnard JG, Singh S, Randolph TW, Carpenter JF. Subvisible particle counting provides a sensitive method of detecting and quantifying aggregation of monoclonal antibody caused by freeze-thawing: insights into the roles of particles in the protein aggregation pathway. *J Pharm Sci 2011*. 2011;100(2):492-503.

6. Twomey A, Less R, Kurata K, Takamatsu H, Aksan A. In Situ Spectroscopic Quantification of Protein-Ice Interactions. *J Phys Chem B*. 2013;117(26):7889-7897. doi:10.1021/jp403267x

7. Cordes AA, Carpenter JF, Randolph TW. Accelerated Stability Studies of Abatacept Formulations: Comparison of Freeze–Thawing- and Agitation-induced Stresses. *J Pharm Sci*. 2012;101(7):2307-2315.

8. Sluzky V, Klibanov AM, Langer R. Mechanism of Insulin Aggregation and Stabilization in Agitated Aqueous-Solutions. *Biotechnol Bioeng 1992*. 1992;40:895-903.

9. Webb S, Cleland JL, Carpenter JF, Randolph TW. Protein aggregation at interfaces formed during lyophilization. *Abstr Pap Am Chem Soc*. 2002;224:163-BIOT.

10. Sethuraman A, Morcone T, Belfort G. Protein aggregation at interfaces. *Protein Sci*. 2004;13:223.

11. Ludwig DB, Carpenter JF, Hamel J-B, Randolph TW. Protein Adsorption and Excipient Effects on Kinetic Stability of Silicone Oil Emulsions. *J Pharm Sci*. 2010;99:1721-1733.

12. Ha E, Wang W, Wang YJ. Peroxide formation in polysorbate 80 and protein stability. *J Pharm Sci*. 2002;91:2252-2264.

13. Wasylaschuk WR, Harmon PA, Wagner G, et al. Evaluation of hydroperoxides in common pharmaceutical excipients. *J Pharm Sci 2007*. 2007;96:106-116.

14. Kerwin BA. Polysorbates 20 and 80 used in the formulation of protein biotherapeutics: structure and degradation pathways. *J Pharm Sci*. 2008;97(8):2924-2935.

15. Chi EY. Protein Aggregation in Aqueous Solution -- Mechanism, Thermodynamics, and Kinetics. 2004.

16. Thirumangalathu R, Krishnan S, Brems DN, Randolph TW, Carpenter JF. Effects of pH, temperature, and sucrose on benzyl alcohol-induced aggregation of recombinant human granulocyte colony stimulating factor. *J Pharm Sci*. 2006;95(7):1480-1497.

17. Joubert MK, Luo Q, Nashed-Samuel Y, Wypych J, Narhi LO. Classification and characterization of therapeutic antibody aggregates. *J Biol Chem*. 2011;286(28):25118-25133. doi:10.1074/jbc.M110.160457

18. Arakawa T, Ejima D, Akuta T. Protein aggregation under high concentration/density state during chromatographic and ultrafiltration processes. *Int J Biol Macromol*. 2017;95:1153-1158. doi:10.1016/j.ijbiomac.2016.11.005

19. Sharma A, Anderson S, Rathore AS. Filter clogging issues in sterile filtration. *Biopharm Int 2008*. 2008;21:53-57.

20. Liu L, Randolph TW, Carpenter JF. Particles shed from syringe filters and their effects on agitation-induced protein aggregation. *J Pharm Sci 2012*. 2012;101:2952-2959.

21. Barnard JG, Kahn D, Cetlin D, Randolph TW, Carpenter JF. Investigations into the fouling mechanism of parvovirus filters during filtration of freeze-thawed mAb drug substance solutions. *J Pharm Sci*. 2014. doi:10.1002/jps.23881

22. Tzannis ST, Hrushesky WJM, Wood PA, Przybycien TM. Irreversible inactivation of interleukin 2 in a pump-based delivery environment. *Proc Natl Acad Sci*. 1996;93:5460-5465.

23. Tyagi AK, Randolph TW, Dong A, Maloney KM, Hiitscherich CJ, Carpenter JF. IgG Particle Formation during Filling Pump Omperation: A Case Study of Heterogeneous Nucleation on Stainless Steel Nanoparticles. *J Pharm Sci*. 2009;98:94-104. doi:10.1002/jps

24. Saller V, Hediger C, Matilainen J, et al. Influence of particle shedding from silicone tubing on antibody stability. *J Pharm Pharmacol*. 2016;70(Iii):675-685. doi:10.1111/jphp.12603

25. Kueltzo LA, Wang W, Randolph TW, Carpenter JF. Effects of Solution Conditions, Processing Parameters,and Container Materials on Aggregation of aMonoclonal Antibody during Freeze–Thawing. *J Pharm Sci*. 2008;9(5):1801-1812. doi:10.1002/jps

26. Kolhe P, Amend E, Singh SK. Impact of freezing on pH of buffered solutions and consequences for monoclonal antibody aggregation. *Biotechnol Prog*. 2010;26(3):727-733.

doi:10.1002/btpr.377

27. Vlieland ND, Nejadnik MR, Gardarsdottir H, et al. The Impact of Inadequate Temperature Storage Conditions on Aggregate and Particle Formation in Drugs Containing Tumor Necrosis Factor-Alpha Inhibitors. *Pharm Res*. 2018;35(2):1-11. doi:10.1007/s11095-017-2341-x

28. Nayak A, Colandene J, Bradford V, Perkins M. Characterization of subvisible particle formation during the filling pump operation of a monoclonal antibody solution. *J Pharm Sci*. 2011;100:4198-4204.

29. Siska C, Harber P, Kerwin BA. Shocking Data on Parcel Shipments of Protein Solutions. *J Pharm Sci*. 2020;109:690-695.

30. Chisholm CF, Soucie KR, Song JS, et al. Immunogenicity of Structurally Perturbed Hen Egg Lysozyme Adsorbed to Silicone Oil Microdroplets in Wild-Type and Transgenic Mouse Models. *J Pharm Sci*. 2017;106(6):1519-1527.

31. Fradkin AH, Carpenter JF, Randolph TW. Immunogenicity of aggregates of recombinant human growth hormone in mouse models. *J Pharm Sci*. 2009;98(9):3247-3264. doi:10.1002/jps.21834

32. Freitag AJ, Shomali M, Michalakis S, et al. Investigation of the immunogenicity of different types of aggregates of a murine monoclonal antibody in mice. *Pharm Res*. 2015;32(2):430-444. doi:10.1007/s11095-014-1472-6

33. Rosenberg AS. Effects of protein aggregates: An immunologic perspective. *AAPS J*. 2006;8(3):E501-E507. doi:10.1208/aapsj080359

34. Jiskoot W, Kijanka G, Randolph TW, et al. Mouse Models for Assessing Protein Immunogenicity: Lessons and Challenges. *J Pharm Sci*. 2016;105(5):1567-1575. doi:10.1016/j.xphs.2016.02.031

35. Roberts CJ. Non-native protein aggregation kinetics. *Biotechnol Bioeng*. 2007;98(5):927-938. doi:10.1002/bit.21627

36. Wang W, Roberts CJ. Protein aggregation – Mechanisms, detection, and control. *Int J Pharm*. 2018;550(1-2):251-268. doi:10.1016/j.ijpharm.2018.08.043

37. Sharma DK, King D, Oma P, Merchant C. Micro-Flow Imaging: Flow Microscopy Applied to Sub-visible Particulate Analysis in Protein Formulations. *AAPS J*. 2010;12(3):455-464. doi:10.1208/s12248-010-9205-1

38. Sharma DK, Oma P, Pollo MJ, Sukumar M. Quantification and Characterization of Subvisible Proteinaceous Particles in Opalescent mAb Formulations Using Micro-Flow Imaging. *J Pharm Sci*. 2010;99:2628-2642.

39. Narhi LO, Corvari V, Ripple DC, et al. Subvisible (2-100 µm) particle analysis during biotherapeutic drug product development: Part 1, considerations and strategy. *J Pharm Sci*. 2015;104(6):1899-1908. doi:10.1002/jps.24437

40. Zölls S, Weinbuch D, Wiggenhorn M, et al. Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments. *AAPS J*. 2013;15(4):1200-1211. doi:10.1208/s12248-013-9522-2

41. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

42. Gambe-Gilbuena A, Shibano Y, Krayukhina E, Torisu T, Uchiyama S. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images. *J Pharm Sci*. 2020;109(1):614-623. doi:10.1016/j.xphs.2019.10.034

43. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. *Adv Neural Inf Process Syst*. 2012:1-9. doi:http://dx.doi.org/10.1016/j.protcy.2014.09.007

44. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118. doi:10.1038/nature21056

45. Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2015;07-12-June:815-823. doi:10.1109/CVPR.2015.7298682

46. Sun Y, Chen Y, Wang X, Tang X. Deep learning face representation by joint identification-verification. *Adv Neural Inf Process Syst*. 2014;3(January):1988-1996.

47. Taigman Y, Yang M, Ranzato M, Wolf L. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *IEEE Conf Comput Vis Pattern Recognit*. January 2014. doi:10.1109/CVPR.2014.220

48. Kiese S, PappenBerger A, Friess W, Mahler H-C. Shaken, Not Stirred: Mechanical Stress Testing of an IgG1 Antibody. *J Pharm Sci*. 2008;97(10):4347-4366. doi:10.1002/jps

49. Ennis RD, Pritchard R, Nakamura C, et al. Glass vials for small volume parenterals: Influence of drug and manufacturing processes on glass delamination. *Pharm Dev Technol*. 2001;6(3):393-405. doi:10.1081/PDT-100002248

50. Bee JS, Randolph TW, Carpenter JF, Bishop SM, Dimitrova MN. Effects of surfaces and leachables on the stability of biopharmaceuticals. *J Pharm Sci*. 2011;100(10):4158-4170. doi:10.1002/jps.22597

51. Gerhardt A, McGraw NR, Schwartz DK, Bee JS, Carpenter JF, Randolph TW. Protein aggregation and particle formation in prefilled glass syringes. *J Pharm Sci*. 2014;103(6):1601-1612. doi:10.1002/jps.23973

52. Kumru OS, Liu J, Ji JA, et al. Compatibility, Physical Stability, and Characterization of an IgG4 Monoclonal Antibody After Dilution into Different Intravenous Administration Bags. *J Pharm Sci*. 2012;101(10):3636-3650. doi:10.1002/jps

53. Krayukhina E, Tsumoto K, Uchiyama S, Fukui K. Effects of syringe material and silicone oil lubrication on the stability of pharmaceutical proteins. *J Pharm Sci*. 2015;104(2):527-535.

doi:10.1002/jps.24184

54. Teska BM, Brake JM, Tronto GS, Carpenter JF. Aggregation and Particle Formation of Therapeutic Proteins in Contact With a Novel Fluoropolymer Surface Versus Siliconized Surfaces: Effects of Agitation in Vials and in Prefilled Syringes. *J Pharm Sci*. 2016;105(7):2053-2065. doi:10.1016/j.xphs.2016.04.015

55. Rosenblatt M. Remarks on a Multivariate Transformation. *Ann Math Stat*. 1952;23(3):470-472. doi:10.1214/aoms/1177729394

56. Scott DW. Multivariate Density Estimation.; 2015.

57. Hermans A, Beyer L, Leibe B. In Defense of the Triplet Loss for Person Re-Identification. 2017. http://arxiv.org/abs/1703.07737.

58. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. 2015:1-14.

59. Dugas C, Bengio Y, Bélisle F, Nadeau C, Garcia R. Incorporating second-order functional knowledge for better option pricing. *Adv Neural Inf Process Syst*. 2001.

60. Kingma DP, Ba JL. Adam: A method for stochastic optimization. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. 2015:1-15.

61. Darling DA. The Kolmogorov-Smirnov , Cramer-von Mises Tests. *Ann Math Stat*. 1957;28(4):823-838.

62. Anderson TW, Darling DA. A Test of Goodness of Fit. *J Am Stat Assoc*. 1954;49(268):765-769. doi:10.1017/CBO9781107415324.004

63. Justel A, Peña D, Zamar R. A multivariate Kolmogorov-Smirnov test of goodness of fit. *Stat Probab Lett*. 1997;35(3):251-259. doi:10.1016/S0167-7152(97)00020-5

64. Hong Y, Li H. Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. *Rev Financ Stud*. 2005;18(1):37-84. doi:10.1093/rfs/hhh006

65. Bhatnagar BS, Bogner RH, Pikal MJ. Protein stability during freezing: Separation of stresses and mechanisms of protein stabilization. *Pharm Dev Technol*. 2007;12(5):505-523. doi:10.1080/10837450701481157

66. Searles JA, Carpenter JF, Randolph TW. The ice nucleation temperature determines the primary drying rate of lyophilization for samples frozen on a temperature-controlled shelf. *J Pharm Sci*. 2001;90(7):860-871. doi:10.1002/jps.1039

67. Sarciaux JM, Mansour S, Hageman MJ, Nail SL. Effects of buffer composition and processing conditions on aggregation of bovine IgG during freeze-drying. *J Pharm Sci*. 1999;88(12):1354-1361. doi:10.1021/js980383n

68. Dong J, Hubel A, Bischof JC, Aksan A. Freezing-induced phase separation and spatial microheterogeneity in protein solutions. *J Phys Chem B*. 2009;113(30):10081-10087.

doi:10.1021/jp809710d

69. Strambini GB, Gabellieri E. Proteins in frozen solutions: Evidence of ice-induced partial unfolding. *Biophys J*. 1996;70(2 I):971-976. doi:10.1016/S0006-3495(96)79640-6

70. Strambini GB, Gonnelli M. Protein stability in ice. *Biophys J*. 2007;92(6):2131-2138. doi:10.1529/biophysj.106.099531

Chapter 5: Machine learning and accelerated stress approaches to differentiate potential causes of aggregation in polyclonal antibody formulations during shipping

Based on a manuscript in press: Witeof AE, Daniels AL, Rea LT, Movafaghi S, Kurtz K, Davis M, Calderon CP, Randolph TW. Machine learning and accelerated stress approaches to differentiate potential causes of aggregation in polyclonal antibody formulations during shipping. Journal of Pharmaceutical Sciences. 2021

5.1 Abstract

Therapeutic protein formulations are distributed through wide, complex supply chains to ensure these drugs reach the growing number of patients they are prescribed to. While it is known that shipping these formulations results in protein aggregation, the mechanism(s) of aggregation during shipping are unknown. This study investigated the mechanism of aggregation during shipping using a recently-developed algorithm for analyzing particle populations using flow imaging microscopy (FIM). This algorithm was used to compare aggregate populations formed during real-time shipping stress with those formed during accelerated shipping stresses mimicking the sporadic high g-force events and continuous low g-force events that may contribute to shipping-induced aggregation. A similar comparison was also performed between aggregates formed in two different secondary packaging types to modulate high g-force stresses. These comparisons revealed that aggregate populations formed during real-time shipping stresses resemble those formed by continuous low g-force events and not those

formed by infrequent high g-force events, suggesting that low g-force events play a dominant role in the mechanisms of aggregation during shipping.

5.2 Introduction

After manufacturing, protein formulations must be shipped through wide, complex supply chains to reach the patients that will ultimately use the formulation. Shipping can expose protein formulations to mechanical stresses that can induce protein aggregation[1–6]. The mechanical stresses the formulations are exposed to can be broken down into two major types: high g-force and low g-force stresses. High g-force stresses are infrequent but severe agitation events such as dropping that can induce cavitation in the formulations[3]. Low g-force stresses consist of less severe but more frequent agitation events such as vibrations on a truck or gentle rocking during normal package handling. Previous studies on shipping-induced aggregation often focus on aggregation induced by high g-force events such as dropping during shipping[6] and strategies for mitigating these extreme stresses[7]. However, the relative contribution of infrequent high g-force events and frequent low g-force events towards shipping-induced aggregation is unknown. If the relative contributions of these stresses could be identified, secondary packaging could be developed to mitigate the most harmful stress on the formulation to minimize aggregation during shipping.

Flow imaging microscopy (FIM)-based aggregate morphology analysis may be useful for probing the relative contributions of low and high g-force events on aggregation during shipping. It has been shown that aggregates formed by different mechanisms exhibit characteristic morphologies when imaged via FIM[8–10]. It is anticipated that aggregates formed due to constant

low g-force events like vibrations will exhibit characteristic morphologies that differ from those formed by infrequent high g-force events like dropping. We recently developed an algorithm that uses machine learning and statistics approaches to analyze the particle morphology distribution of a sample or the types and relative amounts of different aggregate morphologies present in the sample[10]. This approach also allows the user to check if the particle morphology distribution present in one sample is consistent with that present in a separate sample, strongly suggesting that the aggregates were generated by a common mechanism. This approach could be used to compare the aggregate population generated by shipping stress to those made by representative low and high g-force stresses to identify which stress generated aggregate morphologies that best resembled those generated during shipping.

This study used our recently-developed flow imaging microscopy (FIM)-based particle morphology distribution comparison algorithm to compare aggregates formed by shipping to those generated by low g-force vibrations during a shaking stress and high g-force drops during a tumbling stress. Intravenous immunoglobulin (IVIg) formulations were filled in glass vials and packaged in one of two secondary packaging types to modulate high g-force stresses on the formulation. The packaged formulations were then exposed to either real-time shipping, shaking, or tumbling stresses to generate aggregates. The resulting aggregates were then imaged via FIM and analyzed using the algorithm to compare aggregates formed by the real-time and accelerated agitation stresses as well as by the two packaging materials.

5.3 Materials and Methods

5.3.1 Materials

All stress experiments were performed on formulations of intravenous immunoglobulin (IVIg; Gammagard Liquid; Baxter, Deerfield, IL) in 1x phosphate buffered saline (PBS) at pH 7.4 (Gibco, Carlsbad, California). Primary formulation packaging consisted of 5 mL Type 1 borosilicate glass vials (DWK Life Sciences, Milville, NJ) and butyl rubber stoppers (Fisher, Waltham, MA). Both the vials and stoppers were washed with ethanol and water and allowed to air dry before use.

5.3.2 Sample preparation

Stock 100 mg/mL IVIg solutions were diluted to 0.5 mg/mL using PBS. 8.74 mL of this solution were filled into clean 5 mL glass vials. After capping, the filled vials were then placed in one of two secondary packaging types: lab packaging (LAP) in which the vials were bubble wrapped and placed in a plastic pipette tip box and Sealed Air Packaging (SAP) which consisted of custom-fitted rigid foam obtained from Sealed Air Corporation (Charlotte, NC). SAP was designed to mitigate high g-force stresses on the vials relative to LAP. The packaged vials were then exposed to either real-time shipping stress or one of two accelerated model shipping stresses (shaking or tumbling).

Real-time shipping stress: Vials in LAP and SAP were packaged together in a single cardboard box with ice packs to maintain temperature. This box was then shipped back and forth between Boulder, Colorado and Charlotte, NC.

Shaking stress: To simulate low g-force stresses during shipping, vials in LAP or SAP were placed on an orbital shaker and agitated at room temperature for 30 minutes at 700 RPM.

Tumbling stress: To simulated high g-force stresses during shipping, vials in LAP and SAP were placed in a commercial clothing dryer (Roper, Sarasota, FL) and tumbled at room temperature for 5 minutes. During tumbling, the packaging was repeatedly dropped from a height of approximately 18 inches once every second.

5.3.3 Flow Imaging Microscopy

FIM was used to image the aggregate populations generated by each stress. Samples were analyzed using a FlowCam VS (Yokogawa Fluid Imaging Technologies, Inc., Scarborough, ME) fitted with a FC80 FOV flow cell and a 10x objective. 150 µL of each sample was loaded into the instrument, of which 100 µL was analyzed at a flow rate of 0.05 mL/min. Color images of particles larger than 2 µm in diameter were recorded and analyzed.

5.3.4 Particle Comparison Algorithm

FIM images of particles produced in IVIg solutions during shipping and accelerated stress testing were analyzed with an image analysis algorithm described previously[10]. This approach uses ConvNets to compress FIM images into very low-dimensional image representations or embeddings and statistical tools such as nonparametric density estimates and goodness-of-fit hypothesis testing to determine if a sample exhibits embeddings consistent with those present in a different sample. We briefly describe the application of this algorithm here, noting that the

omitted details such as details of how the ConvNet was implemented and trained match those used in the study described in Chapter 4.[10]

FIM images of particles were imported into Python 2.7 and rescaled to 24 x 24 pixels. A ConvNet was trained on FIM images of IVIg aggregates formed in both LAP and SAP as a result of the two accelerated simulated shipping stresses (tumbling or shaking) using a triplet loss approach[11]. The trained network was used to map FIM images onto 2D image representations or embeddings to capture particle information recorded in the raw images. 9,000 FIM images of aggregates generated by each combination of accelerated shipping stress and packaging were used to train this network. The trained network was applied to the remaining (test) images collected from all conditions—including those from the live shipping study that were not used to train the initial ConvNet. The resulting point cloud of 2D embeddings were used to compare particles formed by different stresses (real-time shipping or the accelerated shipping stresses) in the same packaging or different packaging when exposed to the same stress. These comparisons were performed both by visual inspection of the embeddings from these conditions as well goodness-of-fit hypothesis testing. To perform hypothesis testing, the ConvNet was used to compute embeddings of all training images recorded under one of the conditions being compared, which we denote as the baseline condition. The resulting embeddings were then used to construct a kernel density estimate of the probability density function (PDF) of embeddings for the baseline condition. Sets of 20 test images were then subsampled from the images recorded under another condition (denoted as the test condition), converted to 2D embeddings using the trained ConvNet and analyzed using a goodness-of-fit hypothesis test to test the null hypothesis that these embedding sets were consistent with the embedding PDF for the baseline sample. This analysis was performed using a Rosenblatt transform-based goodness-of-fit hypothesis test[10] using a 5% Type I (false positive) error rate to determine critical values for the test. This process was repeated 10,000

times per pair of baseline and test samples and the rejection rate or fraction of these subsamples that were not consistent was recorded as a measure of dissimilarity between the conditions. Samples with very similar particle morphology populations yielded rejection rates close to the 5% Type I error rate of the test whereas samples with more significant differences in particle populations yielded higher rejection rates.

5.4 Results

5.4.1 Flow Imaging Microscopy (FIM)

FIM was used to collect images of protein aggregates generated by real-time and accelerated shipping stresses in both LAP and SAP. Sample FIM images from these samples are shown in

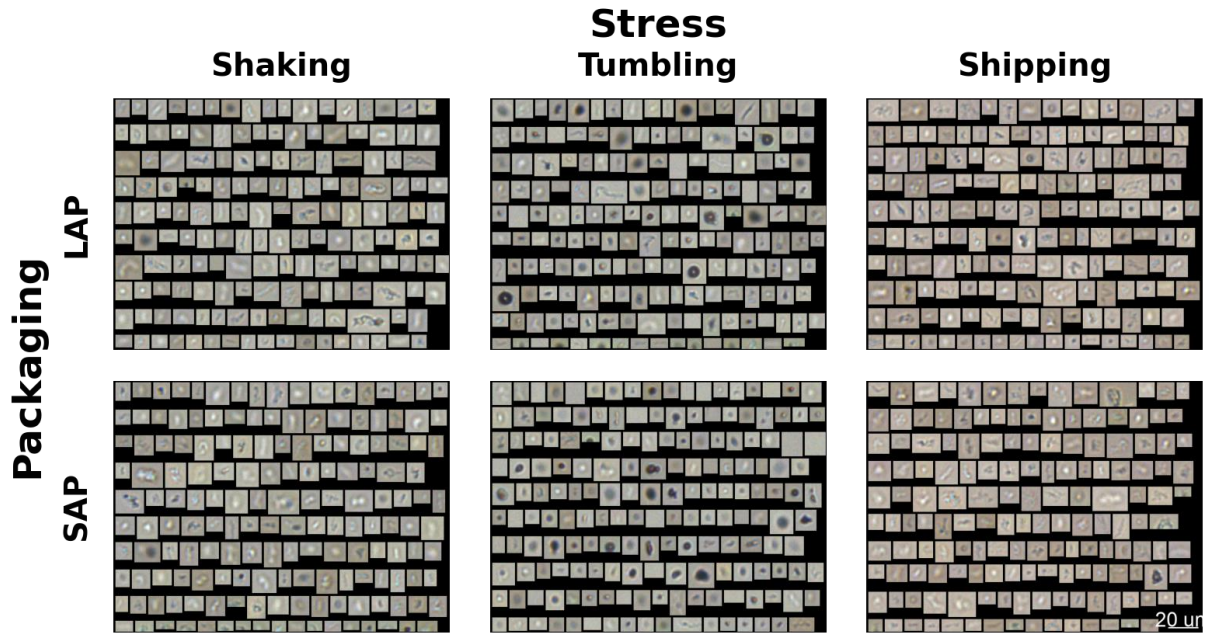

Figure 14: Shipping study FIM images
Randomly selected FIM images of particles from vials packaged in LAP or SAP. Particles were formed via shaking, tumbling, or shipping the vials.

Figure 14. As expected, particles generated by different accelerated stress conditions exhibited visually distinguishable particle morphologies; particles generated by shaking stress generally were small and lightly colored while those generated by tumbling stress were darker on average. Many particles during tumbling qualitatively resembled an air bubble but often with small protein aggregates attached to the surface of the bubble. Particles generated by shipping stress qualitatively resembled the lighter particles generated by shaking rather than the darker particles generated by tumbling.

5.4.2 Particle Comparison Algorithm

Raw FIM images were converted using a ConvNet to 2D embeddings that best captured the particle information in the full image. Figure 15 shows the estimated population density of image embeddings obtained from aggregates generated by shaking, tumbling, and shipping stresses when packaged in either LAP or SAP. The embedding space learned by the network appears to divide FIM image embeddings into two main clusters on the left and right side of the embedding space. These groups delineate particles formed by shaking and tumbling stress: aggregates formed by shaking generally embed in the right cluster while those formed by tumbling generally embed in the left cluster. Figure 15 also shows sample FIM images from each stress condition that maps to specific regions of the embedding space. FIM images that embed in the left cluster typically exhibit dark, circular morphologies while those that embed in the right cluster exhibit lighter, more amorphous morphologies. FIM images that were mapped to the same region of the embedding space in Figure 15 exhibit visually similar particle morphology between different samples. While it is important to note that the image features used to compute image embeddings are not always readily human interpretable[10,12], the visual consistency between FIM

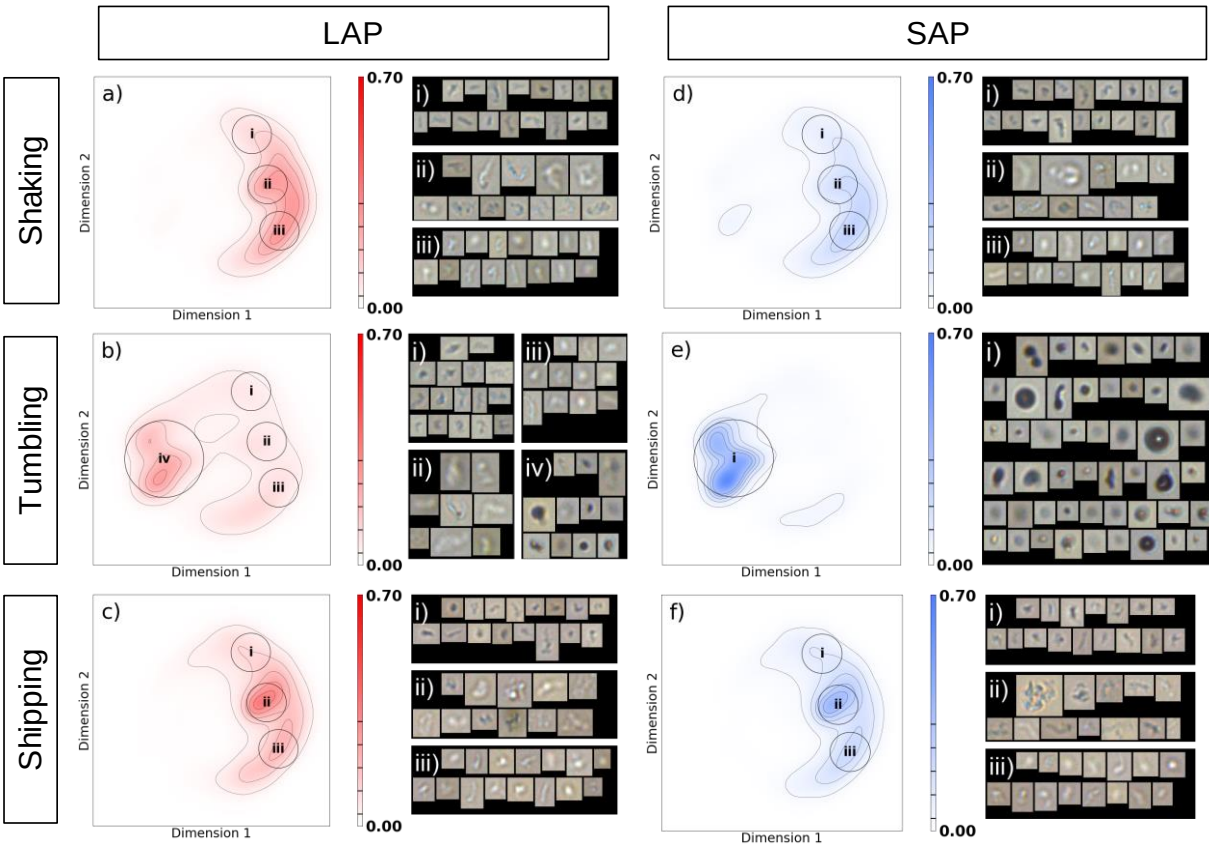

Figure 15: Shipping study embeddings and FIM image collages
Contour plots showing the probability density function of ConvNet-derived FIM image embeddings from IVIg aggregates produced in vials filled to a 1.3 cm^{-1} IAV ratio that had been (a) shaken, (b) tumbled, or (c) shipped in LAP as well as those (d) shaken, (e) tumbled, or (f) shipped in SAP. Dimensions 1 and 2 are arbitrary dimensions learned by the trained ConvNet to differentiate the embeddings of aggregates generated by shaking and tumbling as well as those generated in LAP and SAP. The color at a given region of each plot indicates the fraction of FIM image embeddings for that sample located at that region of the embedding space. The density of embeddings at different regions of the figure can be determined using the color bar associated with each figure. Sample FIM images from each image that embedded in the circled regions of the embedding space are shown in the collages to the right of each figure.

images that are mapped to a single region of the embedding space suggests that these embeddings may be related to visually observable particle morphology features.

The ConvNet trained in this analysis was able to easily differentiate particle morphologies generated by shaking from those arising from application of tumbling stress but was less

effective at differentiating particle morphologies in vials packaged in LAP compared to those packaged in SAP. Particle morphology distributions produced by tumbling were more impacted by packaging than those produced by shaking. Particles formed by tumbling vials in SAP instead of LAP packaging exhibited increased embedding density near the leftmost mode of the PDF, suggesting different mechanisms of aggregation between the two packaging types. In contrast, particles produced by shaking vials packaged in SAP exhibited embeddings relatively similar to those of particles made in LAP packaging and were thus likely formed by similar mechanisms. These results were confirmed using goodness-of-fit hypothesis testing. Sets of 20 randomly selected image of particles produced under shaking stress in SAP packaging were differentiable from those produced in LAP packaging 21% of the time, but 67% of the time when particles were produced under tumbling stress. While particles formed by both stresses were impacted by the packaging, the higher rejection rate for particles made by tumbling suggests a stronger packaging effect for that stress overall.

As shown in Figure 15, the embedding maps of particles produced by shipping stresses in both types of packing materials closely resembled those particle populations generated in accelerated shaking stress studies. Aggregates formed by shipping stress also exhibited visually similar morphologies to those generated by shaking stress (Figure 25, collages). This result was confirmed using goodness-of-fit hypothesis testing. Randomly selected sets containing 20 FIM images of particles recorded in sample packaged in SAP could not be distinguished from those produced in LAP at rates higher than 5%, the Type I error rate of the test.

5.5 Discussion:

This study investigated the role of constant low g-force stresses and sporadic high g-force stresses on aggregates formed during real-time shipping stresses. Shipping exposes protein formulations to low g-force stresses (e.g. shaking[1,13]) and high g-force stresses (e.g. dropping[14,15]) that may trigger aggregation. The ConvNet-generated 2D embeddings of FIM images obtained from shipped samples were most similar to those obtained from shaken samples which were only exposed to continuous, low g-force motions (Figure 25). The particles produced by shaking and shipping stress were also distinct from those produced by tumbling stress that exposed these samples to repeated high g-force events. These results suggest that low g-force events that shipped samples were continuously exposed to such as sample vibrations on a truck were more responsible for aggregation during shipping than infrequent high g-force events such as dropping.

The minimal impact of secondary packaging on particle morphology distributions is consistent with the dominance of low g-force stresses on aggregation during shipping. SAP packaging was designed to attenuate high g-force events on samples[16] but offered little additional protection over LAP against low g-force events. This design resulted in a large impact of packaging on samples exposed to tumbling stress in which these high g-force events dominate the mechanism of aggregation. In contrast, similar aggregate populations were generated in both secondary packaging types during shaking stresses that induced aggregation primarily via low g-force stresses. Similarly, packaging did not influence the particle populations generated during shipping stress. The similar particle populations generated in both packaging types during shipping stress despite the dampened effect of high g-force stresses also suggests that low g-force stresses dominated the mechanisms of aggregation during shipping.

5.6 Acknowledgements

This work was supported by Sealed Air Corporation.

5.7 References

1. Kiese S, PappenBerger A, Friess W, Mahler H-C. Shaken, Not Stirred: Mechanical Stress Testing of an IgG1 Antibody. *J Pharm Sci*. 2008;97(10):4347-4366. doi:10.1002/jps

2. Bee JS, Schwartz DK, Trabelsi S, et al. Production of particles of therapeutic proteins at the air-water interface during compression/dilation cycles. *Soft Matter*. 2012;8(40):10329-10335. doi:10.1039/c2sm26184g

3. Randolph TW, Schiltz E, Sederstrom D, et al. Do Not Drop : Mechanical Shock in Vials Causes Cavitation , Protein Aggregation , and Particle Formation. *J Pharm Sci*. 2015;104:602-611. doi:10.1002/jps.24259

4. Torisu T, Maruno T, Hamaji Y, Ohkubo T, Uchiyama S. Synergistic Effect of Cavitation and Agitation on Protein Aggregation. *J Pharm Sci*. 2017;106(2):521-529. doi:10.1016/j.xphs.2016.10.015

5. Snell JR, Monticello CR, Her C, et al. DEHP Nanodroplets Leached From Polyvinyl Chloride IV Bags Promote Aggregation of IVIG and Activate Complement in Human Serum. *J Pharm Sci*. 2020;109(1):429-442. doi:10.1016/j.xphs.2019.06.015

6. Siska C, Harber P, Kerwin BA. Shocking Data on Parcel Shipments of Protein Solutions. *J Pharm Sci*. 2020;109:690-695.

7. Dill S, Brees K, Stahly A, Cheng E, Carpenter J, Caplan L. Mechanical Shock During Shipping of Medications: The Roles of Packaging and Transportation Vendors. *J Pharm Sci*. 2020;109(1):670-676. doi:10.1016/j.xphs.2019.10.050

8. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

9. Gambe-Gilbuena A, Shibano Y, Krayukhina E, Torisu T, Uchiyama S. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images. *J Pharm Sci*. 2020;109(1):614-623. doi:10.1016/j.xphs.2019.10.034

10. Daniels AL, Calderon CP, Randolph TW. Machine Learning & Statistical Analyses for Extracting and Characterizing “Fingerprints” of Antibody Aggregation at Container Interfaces from Flow Microscopy Images. *Biotechnol Bioeng*. 2020.

11. Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2015;07-12-June:815-823. doi:10.1109/CVPR.2015.7298682

12. Wang H, Wu X, Huang Z, Xing EP. High Frequency Component Helps Explain the Generalization of Convolutional Neural Networks. 2019;(Remark 1):8684-8694. http://arxiv.org/abs/1905.13545.

13. Teska BM, Brake JM, Tronto GS, Carpenter JF. Aggregation and Particle Formation of Therapeutic Proteins in Contact With a Novel Fluoropolymer Surface Versus Siliconized Surfaces: Effects of Agitation in Vials and in Prefilled Syringes. *J Pharm Sci*. 2016;105(7):2053-2065. doi:10.1016/j.xphs.2016.04.015

14. Wu H, Chisholm CF, Puryear M, et al. Container Surfaces Control Initiation of Cavitation and Resulting Particle Formation in Protein Formulations After Application of Mechanical Shock. *J Pharm Sci*. 2020;109(3):1270-1280. doi:10.1016/j.xphs.2019.11.015

15. Movafaghi S, Wu H, Francino Urdániz IM, et al. The Effect of Container Surface Passivation on Aggregation of Intravenous Immunoglobulin Induced by Mechanical Shock. *Biotechnol J*. 2020;15(9):1-9. doi:10.1002/biot.202000096

16. Franc J-P, Michel J-M. *Fundamentals of Cavitation*. Vol 76. Springer science & Business media; 2006.

Chapter 6: Mechanical shock induces fibril formation: effect of mechanical stresses and surface treatments on fibril morphology

Based on a manuscript in preparation: Movafaghi S, Wu H, Daniels AL, Francino Urdániz *IM, Rowe TM, Bull DS, Goodwin A, Randolph TW. Evaluating Amyloid Fibril Formation Caused by Mechanical Shock and the Effect of Stress Type on Fibril Morphology*

6.1 Abstract

The aggregation of amyloid beta and insulin proteins into amyloid fibrils plays a critical role in the progression of diseases such as Alzheimer's disease and insulin-derived amyloidosis. It is hypothesized that severe mechanical shock and the resulting cavitation events may cause these proteins to form fibrils. Flow imaging microscopy (FIM) combined with machine learning methods can be used to monitor changes in aggregate morphology that result from changes in aggregation mechanisms. This study used these FIM-based approaches to determine if changes in aggregate morphologies formed in insulin and amyloid beta formulation following cavitation-inducing mechanical stresses were consistent with fibrillation during these stresses. Formulations of these proteins were exposed to one of two mechanical stresses to induce cavitation and then incubated at either room temperature or 40°C to modulate fibril growth. If cavitation triggers fibrillation, samples exposed to either mechanical stress would contain similar aggregate populations following 40°C incubation. As hypothesized, each protein yielded similar aggregate populations following cavitation and elevated temperature incubation, consistent with

fibril formation following severe mechanical stress and cavitation. These results suggest that cavitation and severe mechanical stress in general may trigger fibril formation and may contribute to fibril-associated human diseases.

6.2 Introduction

Proteins found *in vitro* in therapeutic protein formulations and those found *in vivo* are often exposed to severe mechanical shocks such as dropping or physical trauma. These mechanical stresses have been shown *in vitro* to trigger cavitation or the formation and collapse of gas bubbles in the solution the protein is in[1–3]. The formation and collapse of these bubbles can result in large spikes in local temperature[4,5] as well as the temporary formation and destruction of air-water interfaces[6] both of which may contribute to protein aggregation. It is hypothesized that cavitation and severe mechanical stress in general may cause proteins such as insulin and amyloid beta to form amyloid fibrils or protein aggregates that consist of intermolecular β-sheet structures. These fibrils not only nucleate additional fibril formation[7,8] but are also associated with human diseases such as Alzheimer's disease.

The formation of fibrils during cavitation is hypothesized since many human diseases involving these fibrils are more common following exposure to severe mechanical shock. For example, it has been observed that athletes in sports like American football, boxing, or sledding can exhibit chronic traumatic encephalopathy (CTE) and other neurodegenerative diseases after repeated and/or severe physical trauma from playing their sport[9–12]. While it is unknown if these shocks trigger cavitation in the human brain, it is possible that these mechanical shocks may induce fibrillation of proteins like amyloid beta associated with the progression of these neurodegenerative diseases.

Similar fibrillation in insulin formulations following mechanical shock may also trigger insulin-derived amyloidosis, a rare side effect of insulin injections in which large, insoluble insulin fibrils are formed at the injection site that degrade glucose regulation[13] and may be toxic to the surrounding tissue[14]. Patients often carry insulin formulations with them in pockets or bags which increases the risk of dropping these therapeutics. Cavitation events from dropping these insulin formulations may induce fibril formation. Upon injection, these fibrils would then nucleate the growth of large insulin fibrils characteristic of this condition.

To test if cavitation-inducing mechanical stress can trigger insulin and amyloid beta fibrillation, one could expose formulations of these proteins to cavitation-inducing stresses and test for fibrils in the stressed sample. While it is anticipated that these stresses may generate trace amounts of fibrils, exposing these stressed samples to conditions such as elevated temperature to accelerate additional fibril formation and growth[15] may make a trace fibril concentration detectable. These fibrils can then be detected using standard techniques such as Thioflavin T fluorescence[16].

A more unorthodox strategy to perform this analysis is to use flow imaging microscopy (FIM) combined with morphology analysis algorithms to track changes in aggregate populations resulting from cavitation. FIM instruments capture light microscopy images of all particles present in the sample[17,18]. While fibrils are too small to be imaged using light microscopy, it has been shown that the particle morphologies present in aggregates large enough to be imaged via FIM are sensitive to the mechanisms of aggregation that formed them[19,20]. To perform testing via this approach, one would expose formulations of these proteins to multiple cavitation-inducing mechanical stresses and expose the stressed samples to elevated temperature to promote additional fibrillation and fibril growth. The resulting aggregate populations would then be compared using this FIM image analysis. If these mechanical stresses triggered fibril

formation, each stress should generate a similar aggregate population after incubation due to the common underlying mechanism of aggregation.

The goal of this study was to use flow imaging microscopy (FIM) and particle morphology analysis algorithms to determine if cavitation-inducing stresses induce fibril formation in insulin and amyloid beta formulations. Formulations of these proteins were exposed to sonication and tumbling stresses to induce aggregation. The stressed samples were then incubated at either room temperature or elevated temperature to modulate fibril growth and imaged by FIM. The resulting FIM images were then compared to identify if the two stresses resulted in similar particle populations after elevated temperature incubation—behavior that is consistent with fibril formation.

6.3 Materials and Methods

6.3.1 Materials

Three proteins were used in this analysis. Human insulin powder was obtained from Biogems (Westlake, California) and was formulated in a 15 mM sodium phosphate buffer with 1% sodium azide at pH 3. Amyloid beta 40 (Aβ40) and amyloid beta 42 (Aβ42) powders were obtained from ABclonal Technology (Wobun, MA). Powdered protein samples were dissolved in DMSO solution before use. Working solutions were made from the dissolved protein powders by dilution with 1x phosphate buffered saline (PBS) at pH 7.4 (Gibco, Carlsbad, CA). The final protein concentration was 0.5 mg/mL for insulin formulations and 0.1 mg/mL for amyloid beta formulations.

6.3.2 Stresses

Formulations of insulin, Aβ40 and Aβ42 were prepared by cleaning 2 mL glass vials (Sigma Aldrich, St. Louis, MO) and butyl stoppers (DWK Life Sciences, Germany) with ethanol and water and filling the vials with protein solution. Several of the filled vials were then agitated using either sonication or tumbling stress. Sonication stress consisted of suspending the samples in a sonication bath for one hour. Tumbling stress involved placing these glass vials in a plastic pipette tip box and agitating the box in a commercial clothing dryer (Roper, Sarasota, FL) for 30 minutes. The dryer dropped the samples approximately once per second from a height of approximately 18 inches.

Stressed and control samples were incubated either at room temperature or at 40°C for three days after the initial stress was performed. The elevated temperature incubation was used to promote further formation of fibrils. Three replicates were generated for each combination of protein, stress condition (control, sonication, or tumbling) and incubation temperature.

6.3.3 Flow Imaging Microscopy (FIM)

Flow imaging microscopy (FIM) was used to image protein aggregate populations before and after application of different stress and incubation conditions. Two FlowCams (Yokogawa Fluid Imaging Technologies, Inc., Scarborough, ME) were used for this analysis: a FlowCam VS for Aβ40 and Aβ42 samples and a FlowCam Nano for insulin samples. The FlowCam Nano was required to analyze insulin samples as it was observed that aggregates of this protein were smaller than those for the other proteins and barely large enough to be detected on the FlowCam VS. Both instruments were configured to output color FIM images. FlowCam VS

measurements were performed on 100 µL of samples while those performed using the FlowCam Nano used 200 µL of sample. Ultrapure water was used to clean the flow cell between each sample measurement.

6.3.4 FIM Image Analysis

An algorithm for comparing the morphology of aggregates imaged via FIM[20] was used to compare FIM images from of insulin, Aβ40 and Aβ42 formulations after each combination of mechanical stress and incubation temperature. Briefly, the raw FIM images were exported from the FlowCam and resized to either 27 x 27 pixels for FlowCam VS images or 40 x 40 pixels for FlowCam Nano images. Convolutional neural networks (ConvNets) were trained to learn 2D representations of individual FIM images that clustered images of aggregates formed under a single combination of mechanical stress and incubation temperature together and separated images formed under different conditions. These networks were trained using a triplet loss approach different from that described in Chapter 4 and described in detail in Appendix A. Triplets were mined using a batch all (nonzero) approach[21] in which minibatches consisting of 16 random training FIM images from each samples used in training were constructed and all valid triplets that can be constructed from this minibatch with nonzero triplet loss were used to train the ConvNet. The number of minibatches per epoch was set so that the total number of sampled images in each epoch matched the total number of training images. After each epoch, the triplet loss over a small set of validation images from each sample was computed. The triplet loss over the validation set was used as early stopping criteria and training was stopped after five epochs without improvement to this loss. The model parameters that minimized the triplet loss for the validation set was then used in further analysis.

Two ConvNets were trained for this analysis: one using FIM images of Aβ40 and Aβ42 collected with a FlowCam VS and one using FIM images of insulin collected with the FlowCam Nano. A single network was used to analyze aggregates of both amyloid beta proteins as the proteins were anticipated to yield similar aggregate morphologies. 4,000 training images and 500 validation images from all Aβ42 samples and all but one Aβ40 samples were used to train the ConvNet to analyze amyloid beta aggregates. A reduced number of FIM images from the Aβ40 sample stressed via sonication and incubated at 40°C (1800 training FIM images and 225 validation images) was used due to the limited number of available FIM images from this sample.16,000 training images and 2,000 validation images per stress condition from the insulin samples were used to train the network for that protein. The remaining FIM images from all samples including those from samples not exposed to mechanical stress were analyzed with the appropriate trained network to obtain 2D embeddings. The resulting point cloud of embeddings for each sample were then used to estimate probability density functions (PDFs) for embeddings for that sample using a kernel density estimate. These PDFs were then compared visually between formulations to determine the impact of these stresses on aggregate morphology.

6.4 Results

6.4.1 FIM image analysis of fibril inducing stresses

Flow imaging microscopy (FIM) in combination with a machine learning algorithm for analyzing particle morphology[20] was used to compare the morphologies of large aggregates present in insulin and amyloid beta formulations following mechanical stress and incubation. Figure 16 shows the ConvNet-derived FIM image embeddings from insulin samples. The trained network divided aggregates into three primary clusters. Two of these clusters appeared to be rich in

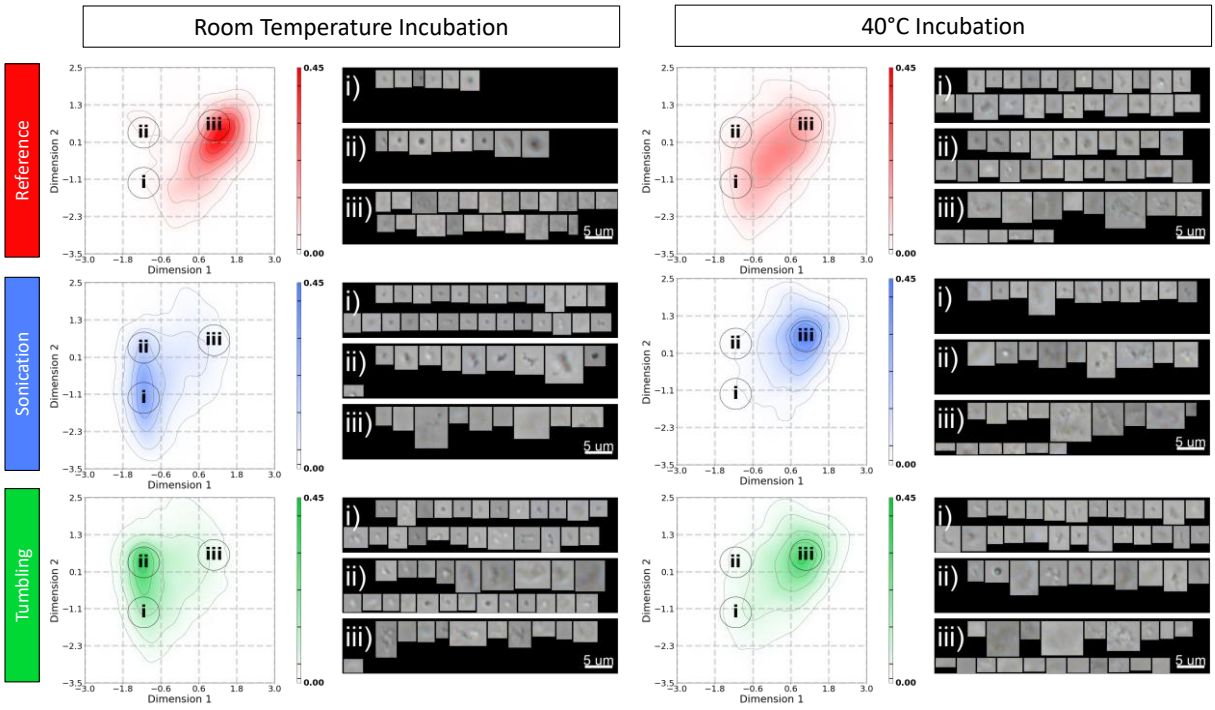

Figure 16: Insulin mechanical shock embeddings and FIM image collages
FIM images and image analysis results for insulin samples after mechanical stress and incubation. Dimensions 1 and 2 are dimensions returned by the trained ConvNet that minimized the triplet loss of these insulin aggregates. Contour colors indicate the fraction of training particles whose embeddings were near that location with darker colors indicating regions containing a higher fraction of particles. Collages show sample FIM images that were mapped to the circled regions of the embedding space. Regions i, ii, and iii correspond to the embedding modes of the sample exposed to sonication and room temperature incubation, tumbling and room temperature incubation and sonication and 40°C incubation respectively.

small, simple aggregates while the third appeared to be rich in aggregates with large, often fuzzy structures. FIM images from samples exposed to either mechanical stress and room temperature incubation (Figure 16, left column, bottom two rows) frequently mapped to clusters in the left side the embedding space--a different location than FIM images from unstressed samples embedded (Figure 16, top row). While the two mechanical stresses yielded different aggregate populations following room temperature incubation, incubating these samples at 40°C instead resulted in similar particle populations from both stresses (Figure 16, right column, bottom two rows). In contrast, unstressed samples were more dispersed throughout the embedding space following 40°C incubation.

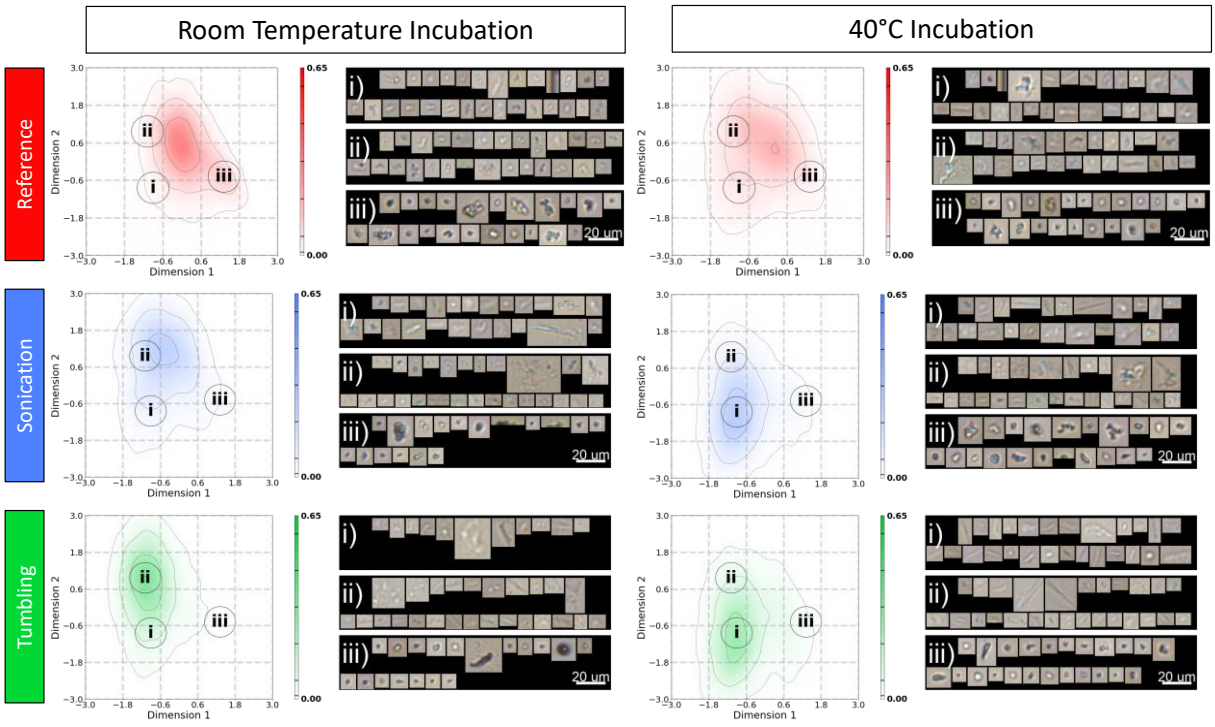

Figure 17: Aβ40 mechanical shock embeddings and FIM image collages
FIM images and image analysis results for Aβ40 samples after mechanical stress and incubation. Contours are interpreted as described for Figure 16. Dimensions 1 and 2 in this figure match those used in Figure 18. Regions i, ii, and iii correspond to the embedding modes for the Aβ40 samples exposed to tumbling stress and 40°C incubation, the Aβ40 samples exposed to tumbling stress and room temperature incubation and the Aβ42 samples exposed to sonication stress and room temperature incubation

Similarly, the combination of mechanical stress with 40°C incubation resulted in similar particle populations in Aβ40 (Figure 17) and Aβ42 (Figure 18) formulations. While sonication and tumbling stresses generated different particle populations in formulations of each protein following room temperature incubation, the combination of either stress with 40°C incubation resulted in a similar particle population for each protein. Both proteins also appear to exhibit a larger fraction of particles in the lower left corner of the embedding space (Figures 17-18, region i) following mechanical stress if incubated at 40°C rather than at room temperature. The raw FIM images that were mapped to this region consisted in part of long, filamentous aggregates that, while too large to be a single fibril, exhibited a fibril-like morphology. While the unstressed

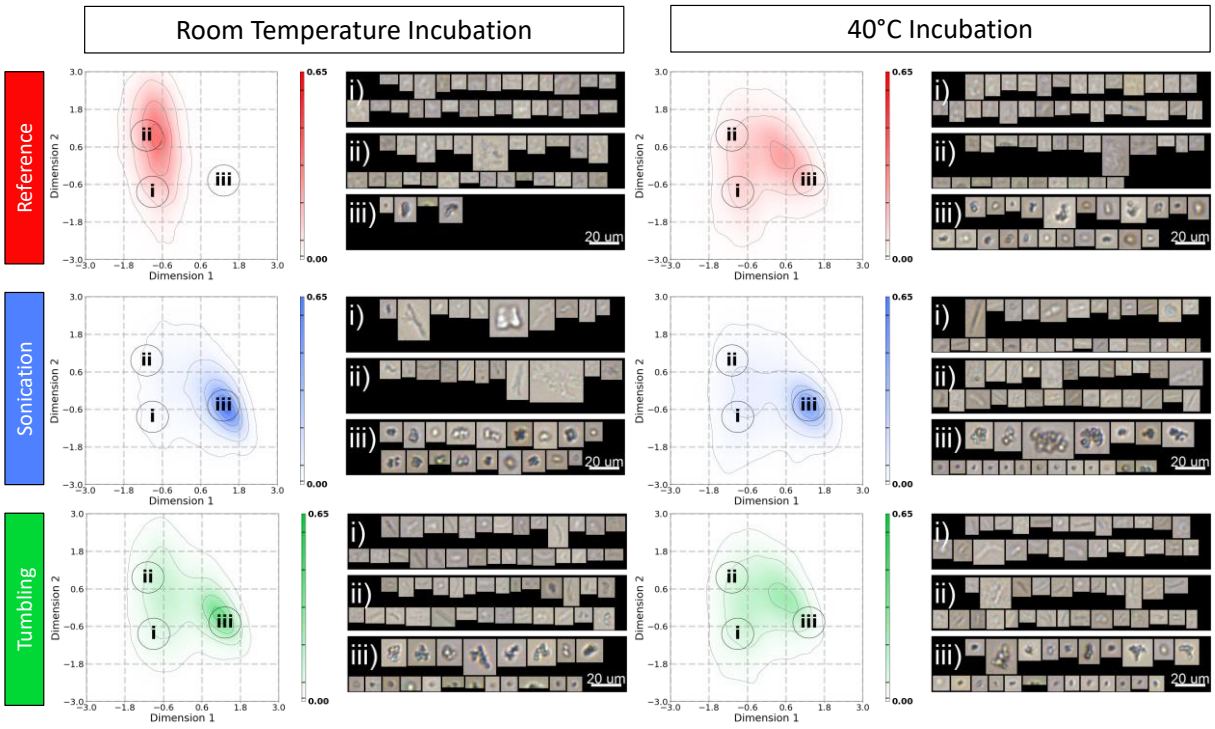

Figure 18: Aβ42 mechanical shock embeddings and FIM image collages
FIM images and image analysis results for Aβ42 samples after mechanical stress and incubation. The format of the subplots matches that described for Figure 17.

reference samples also exhibited particles that embedded in this region, the embedding density in these samples often shifted towards the right of the embedding space after 40°C incubation—the opposite behavior of the stressed samples.

6.5 Discussion

FIM-based strategies for analyzing particle morphologies were used to analyze samples agitated under two cavitation-inducing stresses. These conditions as well as elevated temperature incubation were used to test if insulin, Aβ40 and Aβ42 aggregate morphologies generated by these stresses were consistent with an underlying mechanisms of fibril formation and growth. While FIM-based methods may not appear to be appropriate for this analysis, the

particle morphologies present in FIM images have shown to be sensitive to changes in aggregation mechanisms[19,20,22]. Thus, FIM can be used as an orthogonal technique to other more direct measurements of fibril formation (e.g. ThT fluorescence) to identify changes in aggregate morphology that the formation and growth of fibrils may result in.

Formulations of insulin, Aβ40 and Aβ42 were exposed to two stresses (sonication and tumbling) as well as incubation at elevated temperature to test the ability of these proteins to form fibrils during mechanical stress. If the cavitation during these two stresses induced fibril formation, the combination of mechanical stress and incubation at elevated temperature should trigger the formation of similar particle populations. The ConvNet-derrived FIM image embeddings obtained were consistent with this mechanism. While the initial mechanical stress generated different particle populations in each protein formulation, subsequent incubation at elevated temperatures resulted in more similar aggregate populations. These results are consistent with a mechanism of fibril formation in the stressed samples following incubation at elevated temperature. As similar shifts in embeddings were not observed after incubation at elevated temperature in the unstressed samples, these results also suggest that fibrillation occurred because of the applied stresses. It is important to note that FIM analysis of these aggregates alone does not indicate if the common aggregate morphologies generated by cavitation and incubation is the direct result of fibril formation. Techniques such as imaging flow cytometry[23] may be used in future work to check if these aggregates exhibit a Thioflavin T fluorescence signal characteristic of an underlying amyloid fibril structure.

These results indicate that cavitation-inducing mechanical stresses can trigger fibril formation. These cavitation events may help trigger insulin-derived amyloidosis; since dropping vials is known to induce cavitation[3,24], dropping events from patient mishandling may generate insulin fibrils that can nucleate larger fibrils characteristic with this pathology. However, it is unknown if

trauma experienced by athletes can result in similar cavitation events within the human brain to trigger fibril formation. Thus, it is currently unknown if this mechanism may explain the prevalence of neurodegeneration among athletes after concussions or similar trauma. However, the formation of fibrils during severe mechanical shock *in vitro* raises the possibility that physical trauma *in vivo* may promote the formation of fibrils that are associated with neurodegenerative diseases.

6.6 References

1. Randolph TW, Schiltz E, Sederstrom D, et al. Do Not Drop : Mechanical Shock in Vials Causes Cavitation , Protein Aggregation , and Particle Formation. *J Pharm Sci*. 2015;104:602-611. doi:10.1002/jps.24259

2. Gikanga B, Hui A, Maa Y-F. Mechanistic Investigation on Grinding- induced Subvisible Particle Formation during Mixing and Filling of Monoclonal Antibody Formulations. 2017. doi:10.5731/pdajpst.2017.007732

3. Movafaghi S, Wu H, Francino Urdániz IM, et al. The Effect of Container Surface Passivation on Aggregation of Intravenous Immunoglobulin Induced by Mechanical Shock. *Biotechnol J*. 2020;15(9):1-9. doi:10.1002/biot.202000096

4. Flint EB, Suslick KS. The temperature of cavitation. *Science (80-)*. 1991;253(5026):1397-1399. doi:10.1126/science.253.5026.1397

5. Flannigan DJ, Suslick KS. Plasma formation and temperature measurement during single-bubble cavitation. *Nature*. 2005;434(7029):52-55. doi:10.1038/nature03361

6. Bee JS, Schwartz DK, Trabelsi S, et al. Production of particles of therapeutic proteins at the air-water interface during compression/dilation cycles. *Soft Matter*. 2012;8(40):10329-10335. doi:10.1039/c2sm26184g

7. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem*. 2006;75:333-366. doi:10.1146/annurev.biochem.75.101304.123901

8. Cohen SIA, Linse S, Luheshi LM, et al. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proc Natl Acad Sci U S A*. 2013;110(24):9758-9763. doi:10.1073/pnas.1218402110

9. Walker KR, Tesco G. Molecular mechanisms of cognitive dysfunction following traumatic

brain injury. *Front Aging Neurosci*. 2013;5(JUL):1-25. doi:10.3389/fnagi.2013.00029

10. McKee AC, Stein TD, Nowinski CJ, et al. The spectrum of disease in chronic traumatic encephalopathy. *Brain*. 2013;136(1):43-64. doi:10.1093/brain/aws307

11. Lehman EJ, Hein MJ, Baron SL, Gersic CM. Neurodegenerative causes of death among retired National Football League players. *Neurology*. 2012;79(6):1970-1974. doi:10.1212/01.wnl.0000428873.10254.b7

12. McCradden MD, Cusimano MD. Concussions in sledding sports and the unrecognized “sled head”: A systematic review. *Front Neurol*. 2018;9(SEP):1-8. doi:10.3389/fneur.2018.00772

13. Nagase T, Iwaya K, Iwaki Y, et al. Insulin-derived amyloidosis and poor glycemic control: A case series. *Am J Med*. 2014;127(5):450-454. doi:10.1016/j.amjmed.2013.10.029

14. Iwaya K, Zako T, Fukunaga J, et al. Toxicity of insulin-derived amyloidosis : a case report. 2019:1-6.

15. Kusumoto Y, Lomakin A, Teplow DB, Benedek GB. Temperature dependence of amyloid β-protein fibrillization. *Proc Natl Acad Sci U S A*. 1998;95(21):12277-12282. doi:10.1073/pnas.95.21.12277

16. Hawe A, Sutter M, Jiskoot W. Extrinsic fluorescent dyes as tools for protein characterization. *Pharm Res*. 2008;25(7):1487-1499. doi:10.1007/s11095-007-9516-9

17. Sharma DK, King D, Oma P, Merchant C. Micro-Flow Imaging: Flow Microscopy Applied to Sub-visible Particulate Analysis in Protein Formulations. *AAPS J*. 2010;12(3):455-464. doi:10.1208/s12248-010-9205-1

18. Zölls S, Weinbuch D, Wiggenhorn M, et al. Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments. *AAPS J*. 2013;15(4):1200-1211. doi:10.1208/s12248-013-9522-2

19. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

20. Daniels AL, Calderon CP, Randolph TW. Machine Learning & Statistical Analyses for Extracting and Characterizing “Fingerprints” of Antibody Aggregation at Container Interfaces from Flow Microscopy Images. *Biotechnol Bioeng*. 2020.

21. Hermans A, Beyer L, Leibe B. In Defense of the Triplet Loss for Person Re-Identification. 2017. http://arxiv.org/abs/1703.07737.

22. Gambe-Gilbuena A, Shibano Y, Krayukhina E, Torisu T, Uchiyama S. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images. *J Pharm Sci*. 2020;109(1):614-623. doi:10.1016/j.xphs.2019.10.034

23. Probst C. Characterization of Protein Aggregates, Silicone Oil Droplets, and Protein-Silicone Interactions Using Imaging Flow Cytometry. *J Pharm Sci*. 2019:1-11.

doi:10.1016/j.xphs.2019.05.018

24. Wu H, Chisholm CF, Puryear M, et al. Container Surfaces Control Initiation of Cavitation and Resulting Particle Formation in Protein Formulations After Application of Mechanical Shock. *J Pharm Sci*. 2020;109(3):1270-1280. doi:10.1016/j.xphs.2019.11.015

Chapter 7: Fluorescence microscopy improves sensitivity of machine learning approaches to stress-dependent protein aggregate morphologies

Based on a manuscript in preparation by Daniels AL, Probst C, Calderon CP, Randolph TW. Fluorescence microscopy improves sensitivity of machine learning approaches to stress-dependent protein aggregate morphologies

7.1 Abstract

Machine learning techniques for analyzing protein aggregate morphology information encoded in imaging data have shown promise in identifying protein aggregates formed under different conditions. While many of these algorithms were developed using flow imaging microscopy (FIM) data, in principle these algorithms may also be applied to other imaging modalities that frequently include additional morphology information not accessible from FIM images. In this study, we adapted a previously developed FIM image analysis algorithm to analyze images collected using imaging flow cytometry (IFC), an imaging modality that records fluorescence and light microscopy images for each particle. This algorithm was trained to distinguish between intravenous immunoglobulin (IVIg) aggregates generated by two accelerated stability stresses using both light microscopy images and ProteoStat fluorescence microscopy images. Using both images for each particle resulted in more effective differentiation between aggregate types than was possible using either image type alone. These results demonstrate that the additional particle morphology information found in fluorescence but not light microscopy can be used

during particle morphology analysis to more effectively discriminate between different protein aggregate morphologies.

7.2 Introduction

Protein therapeutics are frequently exposed to conditions in manufacturing[1–3] and shipping[4–6] that can induce the formation of protein aggregates. Manufacturers of protein therapeutics are often interested in monitoring these aggregates and other particle types (e.g. silicone oil microdroplets[7,8]) both to meet current regulations on particle counts (e.g. USP <788> and <789>) and to mitigate the adverse drug responses these particles may be associated with[9,10]. Flow imaging microscopy (FIM)[11,12], a technique used to record light microscopy images of particles in a sample, is an increasingly common technique to perform this particle monitoring. While the images returned by FIM are primarily used to determine particle concentrations, these images can also be analyzed using with machine learning tools such as random forest classifiers[13] or convolutional neural networks (ConvNets)[14,15] to analyze the types of particles present in the sample. The particle morphology available from these light microscopy images (e.g. particle size, shape, color) has previously been used to differentiate between protein aggregates and silicone oil microdroplets[13,16] as well as distinguishing between protein aggregates formed under different stress conditions[14,15,17].

One potential limitation of FIM-based particle morphology analysis is its restriction to particle structure information that can be captured via light microscopy. Other imaging modalities have been developed that capture additional and often orthogonal particle morphology information for each particle. Imaging flow cytometry (IFC) is one such alternative to FIM for this morphology analysis. IFC uses standard fluorescence dyes such as Thioflavin-T (ThT) or ProteoStat to label

the particles in a sample. The resulting fluorescently-labeled particles are then simultaneously imaged via light and fluorescence microscopy in a similar fashion as FIM[18]. These IFC images may be analyzed using algorithms previously developed for FIM image analysis, allowing the user to analyze particle populations using structural information available from both microscopies. Combining these image modalities may allow for better characterization of protein aggregate structure from these algorithms as fluorescent labels are often specific for aggregate structural features on length scales that are not accessible via light microscopy (e.g. Bis-ANS to detect hydrophobic regions[19,20], ThT to detect intermolecular beta-sheet structures[19,21]). Furthermore, as this fluorescence information is captured as an image it is straightforward to generalize existing FIM analysis algorithms to include this additional morphology information. For example, ConvNet-based algorithms can simultaneously analyze light and fluorescence microscopy images to perform morphology analysis.

The focus of this study was to adapt previously developed convolutional neural network (ConvNet) -based algorithm for analyzing light microscopy images captured by flow imaging microscopy (FIM) images to light and fluorescence microscopy images captured by imaging flow cytometry (IFC). Several samples of intravenous immunoglobulin (IVIg) were exposed to one of two stress conditions (freeze-thaw or shaking stress) to form protein aggregates. IFC was then used to capture light and ProteoStat fluorescence images of the aggregates present in each sample. The resulting images were then analyzed via a previously developed ConvNet-based algorithm[15] to compare the particle morphology distribution present in each replicate by these two stresses. This analysis was performed using either the light microscopy images alone, ProteoStat fluorescence images alone, or both images together. These comparisons allowed us to determine if the inclusion of fluorescence information in particle morphology analysis improved the algorithm's ability to differentiate between aggregates formed by the two stresses.

7.3 Methods

7.3.1 Materials

Intravenous immunoglobulin (IVIg; GAMMAGARD LIQUID) was obtained from Takeda International (Lexington, MA). Phosphate buffered saline (PBS) containing 144 mg/L potassium phosphate monobasic, 795 mg/L potassium phosphate dibasic, and 9000 mg/L sodium chloride at pH 7.4 was obtained from Gibco (Waltham, MA). Polypropylene, 2 mL microcentrifuge tubes were from Fisher Scientific (Waltham, MA). Thioflavin T was obtained from Sigma Aldrich (St. Louis, MO). ProteoStat was obtained from Enzo Life Sciences (Farmingdale, NY).

7.3.2 Accelerated Stability Stresses

IVIg aggregates were generated using freeze-thaw and shaking stresses described previously[15]. Briefly, IVIg solution was prepared by centrifuging 100 mg/mL IVIg stock solution at 15,000 x g for 20 minutes to remove aggregates and diluting the supernatant to 0.5 mg/mL with PBS. 1.5 mL aliquots of this stock solution were transferred to water-cleaned 2 mL polypropylene microcentrifuge tubes. Seven of these samples were then exposed to freeze-thaw stress while an additional seven samples were exposed to shaking stress. Freeze-thaw stress consisted of exposing the samples to four freeze-thaw cycles. Each freeze-thaw cycle consisted of a four-minute liquid nitrogen freeze followed by a ten-minute thaw in a 30°C water bath. Shaking stress consisted of taping microcentrifuge tubes horizontally on a plate shaker and agitating the samples at 800 RPM for four hours.

7.3.3 Imaging Flow Cytometry (IFC):

Aggregates present in each sample were imaged using FlowSight (Luminex Corp., Austin, TX) equipped with 488 nm and 785 nm lasers to perform IFC. Before imaging, samples were labeled with two fluorescent dyes: thioflavin-T (ThT) and ProteoStat. During analysis, two 20x magnification grayscale light microscopy images (channels 1 and 9 on the FlowSight) and fluorescence images for emission between 435-505 nm (ThT fluorescence, channel 7 on FlowSight) and emission between 595-642 nm (ProteoStat fluorescence, channel 4 on FlowSight) were collected for each particle in each sample. The raw light and fluorescence microscopy images were then exported from the IDEAS software and rescaled to 30 x 30 pixels before further analysis.

7.3.4 ConvNet Image Analysis

Convolutional neural networks (ConvNets) were trained to distinguish protein aggregates generated from freeze-thaw and shaking stress using subsets of the images collected using FlowSight. The ConvNets were designed to accept a hyperspectral image with each channel of the image consisting of the different images returned by the FlowSight for an individual particle similar to how the red, green, and blue channels of a color image are fed to ConvNets for simultaneous analysis. These networks were designed to return a two-dimensional image representation or embedding describing the information encoded in the hyperspectral image. These representations were optimized using a triplet loss strategy[15]. This setup allowed fluorescence information to be analyzed with the light microscopy images with minimal modifications to the algorithm from that described in chapter 4[15].

Three separate ConvNets were trained to differentiate images of protein aggregates generated by freeze-thaw and shaking stresses. One of these networks was trained using only the two light microscopy images for each image. A second network was trained using only the ProteoStat (channel 4) fluorescence images and a third network was trained using both the light microscopy and ProteoStat fluorescence images. ThT (channel 7) fluorescence images were excluded from these networks as it was observed that these images did not exhibit a fluorescence signal. Each of these networks were trained on a single set of 5,600 particles generated under each stress condition from all replicates with the remaining images excluded from network training used as test data. The trained networks were then applied to all particles to compute embeddings for the test images from each replicate using each network. The resulting embeddings were then used to construct probability density functions (PDFs) of the embeddings from each replicate using a kernel density estimate[22]. These PDFs were then compared visually to compare the embeddings yielded by different replicates of each stress as well as by the three neural networks.

7.4 Results:

7.4.1 Accelerated Stability Stresses

Figure 19 shows sample light microscopy and ProteoStat fluorescence images generated by each stress condition for six randomly selected replicates per stress. While the size of the aggregates in each image is small, the bright halo around each particle and often the ProteoStat fluorescence signal indicated that a particle was present in each image. However, the small size of these images made it difficult to distinguish particles generated by these two stresses by eye.

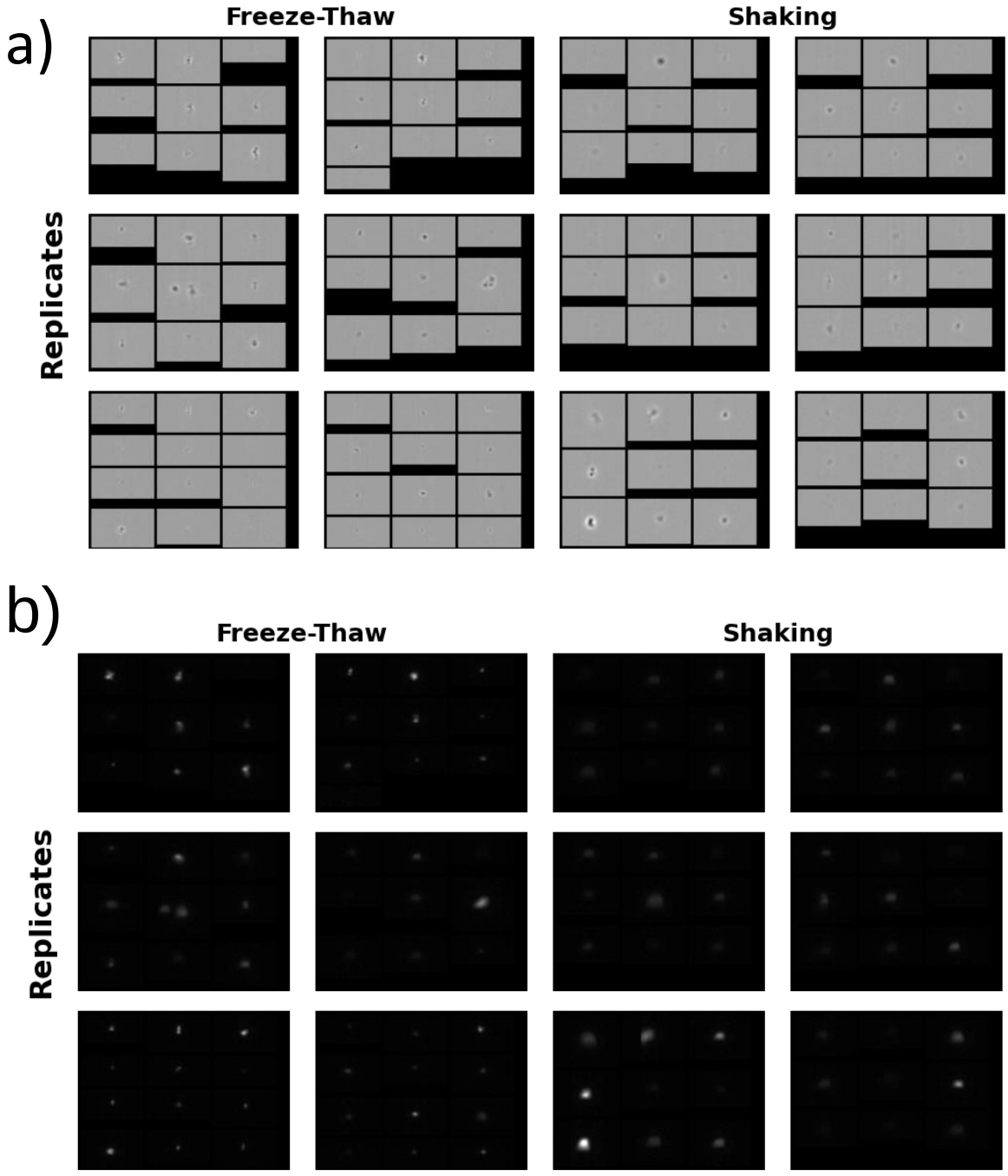

Figure 19: Sample IFC light and fluorescence microscopy images
(a) Sample light microscopy images of aggregates generated by freeze-thaw and shaking stress from each replicate obtained from IFC. (b) Fluorescence microscopy images taken at the emission frequency for ProteoStat for the particles imaged in (a). Each collage contains images from a single replicate with the excluded replicate chosen at random.

7.4.2 ConvNet analysis

Figure 20 shows contour plots of the embedding PDFs for the training data from ConvNets trained on light microscopy images, fluorescence microscopy images, or both images. While all three ConvNets were successfully able to differentiate between particles generated by shaking and freeze-thaw stresses, the tightness of the embedding clustering within each stress as indicated by the height of the embedding PDF's mode varied depending on the images used to compute the embeddings. The embeddings based on fluorescence microscopy images alone exhibited the most diffuse embeddings while those computed from both light and fluorescence images exhibited the tightest embedding clustering.

While aggregates generated by each stress were distinguishable by each network, the embeddings yielded by each of the seven replicates per stress also exhibited different embedding distributions. Figure 21 shows the embedding distributions yielded by the three ConvNets for test images from six randomly selected freeze-thaw replicates. Embedding PDFs are not shown for the individual shaking replicates as not enough images were available to construct accurate PDF estimates for most of the replicates. Each of the three networks identified some variability in the particle embeddings yielded by each replicate. However, the magnitude of this replicate-to-replicate variability depended on the images used to compute embeddings. As was observed in Figure 20, the embeddings yielded by the ConvNet trained on both images yielded the tightest clustering between replicates while those yielded by the network trained only on fluorescence information yielded the most variable embeddings between replicates.

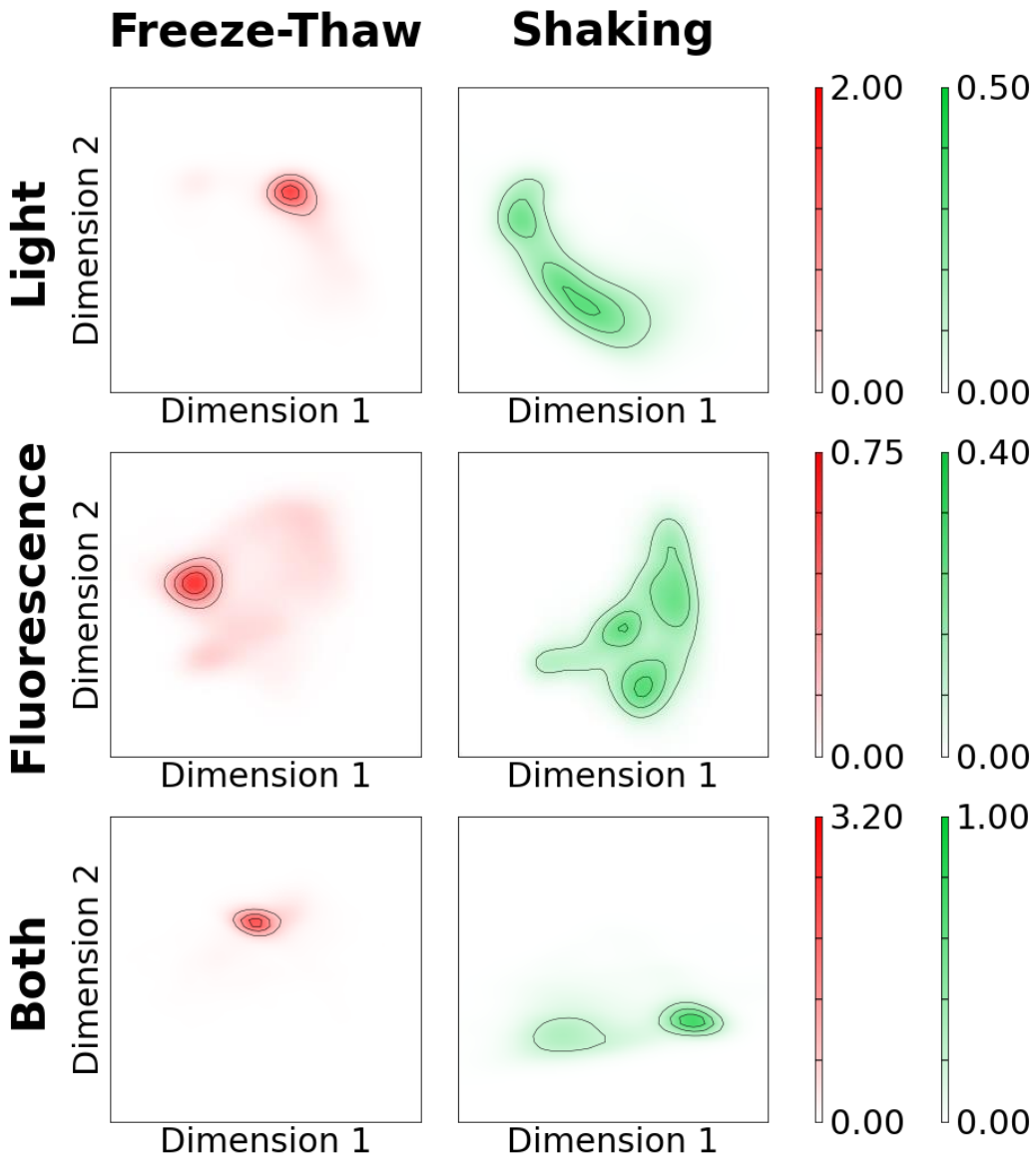

Figure 20: IFC image embeddings from different stresses
Contour plots showing IFC image embeddings from samples exposed to freeze-thaw or shaking stress. Column labels indicate the stress used to generate the aggregates represented by the embeddings while the row labels indicate whether light microscopy images, ProteoStat fluorescence microscopy images, or both were used to compute the embeddings. Contour plots indicate the relative fraction of image embeddings found at a given location of the embedding space with darker colors indicating a greater percentage of image embeddings at that location. The color bar matching the color of the figure can be used to identify the exact fraction of embeddings at a given location. Contour lines were added to aid the viewer visualizing regions of low embedding density and do not have a physical meaning.

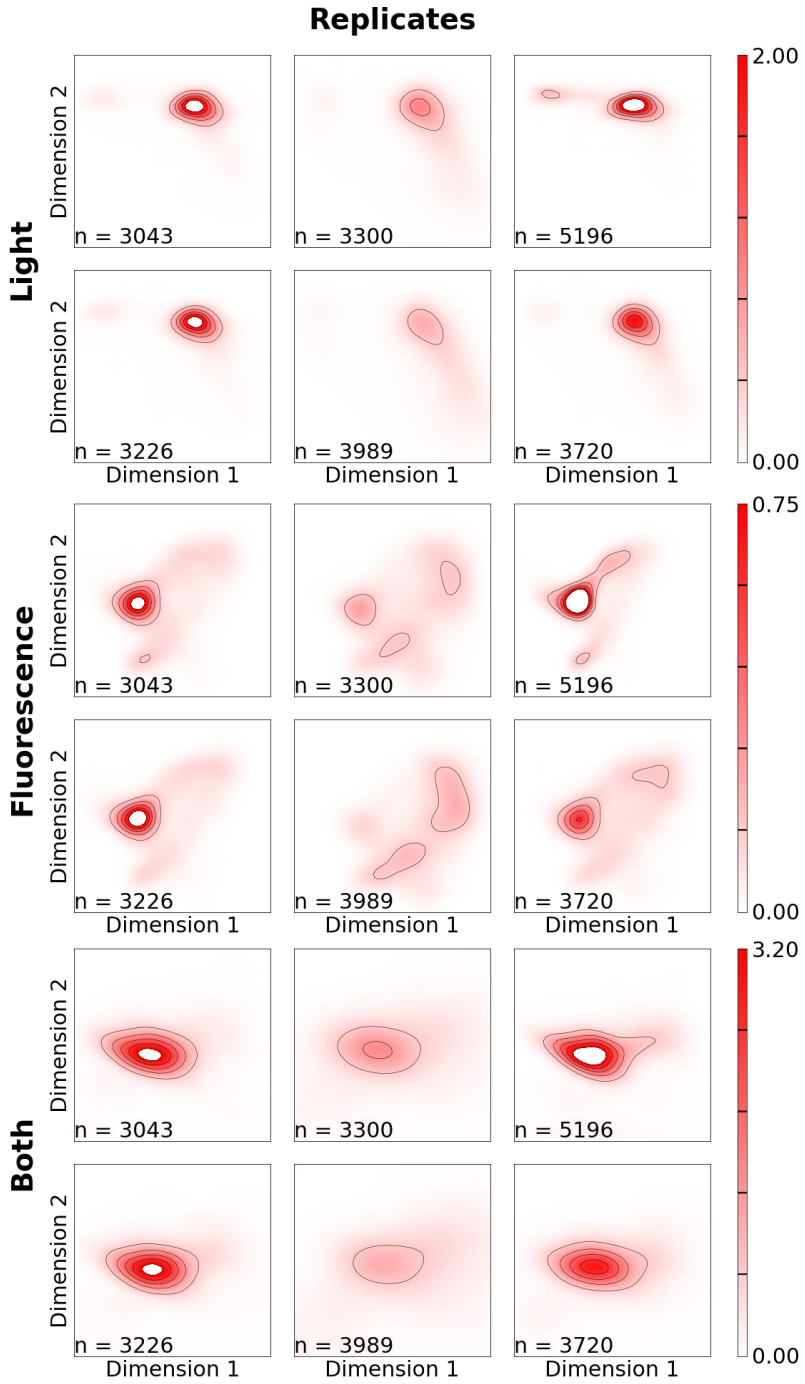

Figure 21: IFC image embeddings from different freeze-thaw replicates
Contour plots showing IFC image embeddings six randomly-selected replicates of freeze-thaw. Row labels indicate the IFC images that were used to compute the embeddings plotted in each pair of rows. Plots in the same position in these row pairs were taken from the same replicate. The number of particle embeddings used to construct each plot are shown in the bottom left corner of the figure. Contours are interpreted as described for Figure 20

7.5 Discussion

This study investigated applying machine learning algorithms developed for FIM data to analyze protein aggregate morphology on images collected from IFC. IFC was used to image particles generated by freeze-thaw and shaking stresses that were previously shown to generate distinguishable particle morphologies that could be detected using FIM and ConvNet-based algorithms[15]. Applying these ConvNet algorithms to images obtained from IFC yielded a similar conclusion; protein aggregates generated by these different stresses did result in distinct particle morphologies (Figure 20). However, the different replicates of these stresses also yielded slightly different embedding distributions (Figure 21). The apparent variability in embeddings between replicates depended on whether light microscopy images, ProteoStat fluorescence microscopy images, or both were used to compute embeddings. Using both light and fluorescence microscopy images resulted in the least variability in embeddings yielded by freeze-thaw replicates and the tightest overall embedding clustering. This improvement over the embeddings yielded by either image type alone suggests that fluorescence microscopy contains particle morphology information that is not encoded in light microscopy information and vice versa. As a result, allowing the ConvNet to analyze particle morphology information from both microscopies results in improved discrimination between the two aggregate types than was possible using light microscopy images alone as is done in typical FIM image analysis.

This study demonstrates how existing particle morphology analysis algorithms can be adapted to IFC and how the additional orthogonal morphology information available through fluorescence microscopy allows for better differentiation between particle types. It is anticipated that these algorithms can also be similarly adapted to analyze imaging data collected by other imaging modalities such as background membrane imaging[23] and holographic video microscopy[24,25]. While many of these modalities still ultimately use light microscopy to capture

particle information and thus capture similar particle morphology information as FIM, the modifications to how particles are imaged (e.g. removing solvent for more effective imaging of translucent particles in background membrane imaging[26]) may still improve the performance of these particle morphology analysis algorithms.

7.6 References

1. Bee JS, Randolph TW, Carpenter JF, Bishop SM, Dimitrova MN. Effects of surfaces and leachables on the stability of biopharmaceuticals. *J Pharm Sci*. 2011;100(10):4158-4170. doi:10.1002/jps.22597

2. Kueltzo LA, Wang W, Randolph TW, Carpenter JF. Effects of Solution Conditions, Processing Parameters,and Container Materials on Aggregation of aMonoclonal Antibody during Freeze–Thawing. *J Pharm Sci*. 2008;9(5):1801-1812. doi:10.1002/jps

3. Her C, Carpenter JF. Effects of Tubing Type, Formulation, and Postpumping Agitation on Nanoparticle and Microparticle Formation in Intravenous Immunoglobulin Solutions Processed With a Peristaltic Filling Pump. *J Pharm Sci*. 2020;109(1):739-749. doi:10.1016/j.xphs.2019.05.013

4. Kiese S, PappenBerger A, Friess W, Mahler H-C. Shaken, Not Stirred: Mechanical Stress Testing of an IgG1 Antibody. *J Pharm Sci*. 2008;97(10):4347-4366. doi:10.1002/jps

5. Bee JS, Schwartz DK, Trabelsi S, et al. Production of particles of therapeutic proteins at the air-water interface during compression/dilation cycles. *Soft Matter*. 2012;8(40):10329-10335. doi:10.1039/c2sm26184g

6. Siska C, Harber P, Kerwin BA. Shocking Data on Parcel Shipments of Protein Solutions. *J Pharm Sci*. 2020;109:690-695.

7. Gerhardt A, McGraw NR, Schwartz DK, Bee JS, Carpenter JF, Randolph TW. Protein aggregation and particle formation in prefilled glass syringes. *J Pharm Sci*. 2014;103(6):1601-1612. doi:10.1002/jps.23973

8. Chisholm CF, Baker AE, Soucie KR, Torres RM, Carpenter JF, Randolph TW. Silicone Oil Microdroplets Can Induce Antibody Responses Against Recombinant Murine Growth Hormone in Mice. *J Pharm Sci*. 2016;105(5):1623-1632. doi:10.1016/j.xphs.2016.02.019

9. Rosenberg AS. Effects of protein aggregates: An immunologic perspective. *AAPS J*. 2006;8(3):E501-E507. doi:10.1208/aapsj080359

10. Kotarek J, Stuart C, De Paoli SH, et al. Subvisible Particle Content, Formulation, and Dose of an Erythropoietin Peptide Mimetic Product Are Associated with Severe Adverse Postmarketing Events. *J Pharm Sci*. 2016;105(3):1023-1027. doi:10.1016/S0022-3549(15)00180-X

11. Sharma DK, King D, Oma P, Merchant C. Micro-Flow Imaging: Flow Microscopy Applied to Sub-visible Particulate Analysis in Protein Formulations. *AAPS J*. 2010;12(3):455-464. doi:10.1208/s12248-010-9205-1

12. Zölls S, Weinbuch D, Wiggenhorn M, et al. Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments. *AAPS J*. 2013;15(4):1200-1211. doi:10.1208/s12248-013-9522-2

13. Saggu M, Patel AR, Koulis T. A Random Forest Approach for Counting Silicone Oil Droplets and Protein Particles in Antibody Formulations Using Flow Microscopy. *Pharm Res*. 2017;34(2):479-491. doi:10.1007/s11095-016-2079-x

14. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

15. Daniels AL, Calderon CP, Randolph TW. Machine Learning & Statistical Analyses for Extracting and Characterizing “Fingerprints” of Antibody Aggregation at Container Interfaces from Flow Microscopy Images. *Biotechnol Bioeng*. 2020.

16. Strehl R, Rombach-Riegraf V, Diez M, et al. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: A novel multiparametric image filter for sub-visible particles in microflow imaging analysis. *Pharm Res*. 2012;29(2):594-602. doi:10.1007/s11095-011-0590-7

17. Gambe-Gilbuena A, Shibano Y, Krayukhina E, Torisu T, Uchiyama S. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images. *J Pharm Sci*. 2020;109(1):614-623. doi:10.1016/j.xphs.2019.10.034

18. Probst C. Characterization of Protein Aggregates, Silicone Oil Droplets, and Protein-Silicone Interactions Using Imaging Flow Cytometry. *J Pharm Sci*. 2019:1-11. doi:10.1016/j.xphs.2019.05.018

19. Hawe A, Sutter M, Jiskoot W. Extrinsic fluorescent dyes as tools for protein characterization. *Pharm Res*. 2008;25(7):1487-1499. doi:10.1007/s11095-007-9516-9

20. Wu H, Randolph TW. Rapid Quantification of Protein Particles in High-Concentration Antibody Formulations. *J Pharm Sci*. 2019;108(3):1110-1116. doi:10.1016/j.xphs.2018.10.021

21. Krebs MRH, Bromley EHC, Donald AM. The binding of thioflavin-T to amyloid fibrils: Localisation and implications. *J Struct Biol*. 2005;149(1):30-37. doi:10.1016/j.jsb.2004.08.002

22. Scott DW. *Multivariate Density Estimation*.; 2015.

23. Helbig C, Ammann G, Menzen T, Friess W, Wuchner K, Hawe A. Backgrounded Membrane Imaging (BMI) for High-Throughput Characterization of Subvisible Particles During Biopharmaceutical Drug Product Development. *J Pharm Sci*. 2019:1-13. doi:10.1016/j.xphs.2019.03.024

24. Wang C, Zhong X, Ruffner DB, et al. Holographic Characterization of Protein Aggregates. *J Pharm Sci*. 2016;105(3):1074-1085. doi:10.1016/j.xphs.2015.12.018

25. Kasimbeg PNO, Cheong FC, Ruffner DB, Blusewicz JM, Philips LA. Holographic Characterization of Protein Aggregates in the Presence of Silicone Oil and Surfactants. *J Pharm Sci*. 2018;108(1):155-161. doi:10.1016/j.xphs.2018.10.002

26. Zölls S, Gregoritza M, Tantipolphan R, et al. How Subvisible Particles Become Invisible — Relevance of the Refractive Index for Protein Particle Analysis. *Pharm Biotechnol*. 2013;102(5):1434-1446. doi:10.1002/jps

Chapter 8: Flow imaging microscopy and machine learning approaches to diagnose bloodstream infections

Based on a manuscript in preparation by Daniels AL, Calderon CP, Randolph TW. Flow imaging microscopy and machine learning approaches to diagnose bloodstream infections

8.1 Abstract

Strategies for rapid, sensitive detection of bloodstream infections (BSIs) are critical for diagnosing sepsis, a severe and potentially fatal immune response to an infection. Flow imaging microscopy (FIM) combined with convolutional neural networks (ConvNets) may be used to image and detect pathogenic cells in a blood sample in a rapid, label-free fashion to diagnose BSIs. As proof-of-concept of this approach, this study demonstrates the use of ConvNets to distinguish between FIM images of blood cells of bacteria and platelets—the blood cell type that is most visually similar to bacterial cells—as well as to distinguish between eight species of bacteria. The trained networks were able to identify new FIM images of these cell types with reasonable single-image accuracy. While further improvements in network accuracy and pathogen scope are necessary for this approach to be useful in practice, these results suggest that FIM-based strategies for BSI diagnosis are feasible.

8.2 Introduction

Sepsis is a severe immune response to an infection that often results in tissue and organ damage as well as potentially patient fatalities[1,2]. At least one million sepsis cases occur annually in the U.S[3] which are associated with a 25-30% mortality rate[4]. Sepsis is particularly common in neonates with 36% of all neonates exhibiting at least one sepsis episode during their initial hospitalization after birth[5] and is associated with a 50% mortality rate[6]. Early treatment of sepsis has been found to improve patient outcomes[7]. Rapid and effective techniques to diagnose sepsis are therefore critical in minimizing patient fatalities and improving patient outcomes to this condition.

Sepsis is often diagnosed using strategies to detect bloodstream infections (BSI) or a direct microbial infection of the patient's bloodstream. Diagnosing these infections requires techniques to detect pathogenic bacterial and yeast cells in a patient's blood. A core challenge with this diagnosis is the extremely low concentration of pathogen cells in patients with sepsis symptoms—patients may exhibit symptoms with as few as 10 colony forming units (CFU) / mL of the causative agent in their bloodstream[8]. This detection is further complicated by the billions of cells per milliliter of human blood consisting of red blood cells, white blood cells and platelets[9]. As a result, techniques for detecting BSIs need high sensitivity to detect trace pathogen concentration against a high background concentration of cells.

Blood culture is the current standard technique for diagnosing BSI-induced sepsis. This protocol involves drawing a 1-10 mL blood sample from a patient and culturing the blood sample in media preferential for bacterial growth. If an organism grows biochemical techniques are then used to identify the species of that organism. Despite advancements in automated approaches[10,11] blood culture is a time-consuming diagnosis strategy that often requires on the

order of days to identify an organism in a blood sample[11,12]. The delayed diagnosis and treatment not only result in worse clinical outcomes but, as patients will often be prescribed broad spectrum antibiotics while blood culture is performed, may result in unnecessary antibiotics use. This approach is also restricted to species that can grow on media for detection and identification--behavior that is not guaranteed for clinical bacteria and yeast strains. Large blood cultures drawn from the patient to are often required for an accurate diagnosis to compensate for both the inherently low pathogen concentration in blood and the difficulty of growing an isolated pathogen. The large blood volume requirement makes it difficult to use blood cultures to diagnose BSIs in neonates as their low body weight greatly limits volume of blood that can be drawn to perform blood cultures.

The problems with blood culture have prompted the development of alternative strategies to detect BSIs for purposes of sepsis diagnosis. One of the most prominent alternative strategies that has been investigated is the use of polymerase chain reactions (PCR) to detect pathogenic DNA in a patient's blood sample[13–15]. While these PCR reactions offer diagnosis times on the order of hours instead of days, these tests are frequently only sensitive to a small albeit usually relevant selection of pathogens as the primers used in the reaction are typically specific to a small set of organisms. As PCR-based methods are not label-free, organisms whose DNA are not amplified by a given PCR scheme will not be detected. This labelling strategy can also result in a false positive diagnosis if DNA from dead pathogens are present in a blood sample[9].

One potentially useful technique for diagnosing sepsis is flow imaging microscopy (FIM), a technique that uses a combination of microfluidics and light microscopy to image particles in a liquid sample larger than approximately 1 μm in diameter in a high-throughput fashion[16,17]. Like images yielded by traditional light microscopy, images obtained from FIM are rich in morphological information about the particles in the sample. This morphology information can

potentially be used to identify different cell types in a sample. For example, this morphology information has previously been used to distinguish between viable and unviable cells[18]. This morphology information may similarly be useful in distinguishing between FIM images of common blood cells and different pathogenic cells in a blood sample to diagnose BSIs. This approach would allow clinicians to identify pathogen cells in a blood sample without the use of culturing or organism-specific labelling that can limit the sensitivity of the diagnosis to specific pathogens. This approach also offers high theoretical sensitivity to pathogens as even a single imaged pathogen may be enough to diagnose an infection. This sensitivity may reduce the blood volume that is needed to obtain an accurate diagnosis which is ideal for diagnosing neonatal sepsis.

Techniques for extracting and analyzing morphology information in FIM images to determine cell identity are critical for a FIM-based sepsis diagnosis to be successful. Convolutional neural networks (ConvNets) have shown promise for this analysis due to their performance in general computer vision tasks[19,20], medical image analysis tasks such as skin cancer diagnosis[21], and FIM image analysis for protein aggregate analysis[22,23]. ConvNets may be able to be trained to recognize FIM images of common blood cells and different species of pathogens which could be used to diagnose BSIs.

This study investigated using convolutional neural networks (ConvNets) to differentiate between flow imaging microscopy (FIM) images of blood cells and different species of bacteria as proof-of-concept of a FIM-based sepsis diagnosis. FIM images of platelets and eight species of bacteria were collected and used to train ConvNets to distinguish FIM images of bacteria from those of platelets as well as to distinguish between different species of bacteria. If these networks can reliably detect each cell type, FIM images of cells in a patient's blood sample can

be analyzed using these networks to identify any pathogenic cells it contains, indicating that a patient does or does not have a BSI.

8.3 Materials and Methods

8.3.1 Materials

Several species of bacteria were grown and imaged and analyzed in this study. *Escherichia coli*, *Serratia marcescens*, and *Lactococcus lactis* cultures were obtained from Microbiologics (St. Cloud, MN) and were derived from ATCC stocks of these organisms (Manassas, VA). Cultures of *Enterococcus faecalis*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii* strains isolated from local clinics were also prepared. Murine blood was taken from female BALB/c mice. These blood draws were performed under the University of Colorado Institute of Animal Care and Use Committee (IACUC) protocol #2318. 2 mL Becton Dickinson Vacutainers with EDTA (Franklin Lakes, NJ) were used in all blood collections. Difco nutrient broth, and brain-heart infusion broth media were prepared using premixed powders obtained from Becton Dickinson (Franklin Lakes, NJ). Lysogeny Broth (LB) and Cation-Adjusted Mueller-Hinton Broth (CAMHB) media were both prepared from reagent-grade chemicals. All media solutions were autoclaved prior to use. 1x phosphate buffered saline (PBS) at pH 7.4 and Dulbecco's Modified Eagle's Media (DMEM) were obtained from Gibco (Carlsbad, California). Ficoll-Paque Premium solution was obtained from GE Healthcare (Chicago, IL). Cargille (Cedar Grove, NJ) immersion oil type 37 was used in FIM measurements.

8.3.2 Bacteria culture preparation

Cultures of each organism in appropriate growth media were grown overnight. *Escherichia coli*, *Serratia marcescens and Lactococcus lactis* cultures were grown in liquid LB, Difco nutrient broth, and brain-heart infusion broth media, respectively. All other cultures were grown in liquid CAMHB media. Cultures were grown in a 37°C incubator with agitation.

8.3.3 Blood sample preparation

1 mL aliquots of blood pooled from several mice after sacrifice via cardiac puncture were obtained and filled into a 2 mL EDTA Vacutainer tube and inverted per manufacturer's instructions. This yielded blood samples with twice the recommended EDTA concentration for human blood which was empirically found to minimize coagulation. These samples were then kept on ice prior to further analysis.

Blood samples contain three major cellular components: red blood cells, white blood cells, and platelets. It was desired to train the algorithm against images of platelets alone instead of a mixture of all three cell types as platelets were anticipated to be more difficult to distinguish from bacterial cells than other blood cells. Platelets exhibit closer size and morphology to the bacterial cells in a sample and, unlike the much larger red and white blood cells, are difficult to remove experimentally[24] or digitally (e.g. filtering out images of large particles collected during FIM). An AcouWash (AcouSort, Lund, Sweden) was used to process the blood samples to obtain concentrated platelet samples for use in algorithm training. This instrument flows separation media sheathed with sample volume through a microfluidics channel and uses standing ultrasound waves to move large cells from the sample to the separation media.[25] The

stream is then separated to yield two outlets enriched in either large or small cells. To perform this separation, blood samples were diluted 1:10 with PBS and flown through the AcouWash along with 70% Ficoll-Paque solution in PBS as the separation media. Sample and media flow rates were 50 and 100 µL/min, respectively. Red blood cell separation was performed using 2041 kHz ultrasound waves. The platelet-rich eluent outlet was set to a 110 µL/min flow rate while the red and white blood cell-rich sample outlet was set to a 40 µL/min flow rate. Note that the final dilution of the platelet-rich outlet was approximately double that of the sample inlet (~1:20) due to the higher flow rate to this outlet. The eluent outlet at these settings were found to contain 99% of the platelets and 1% of the red blood cells present in the initial blood sample.

8.3.4 Flow Imaging Microscopy

Flow imaging microscopy was used to collect images of blood, purified platelets, and bacterial cells. FIM was performed using a FlowCam Nano (Yokogawa Fluid Imaging Technologies, Inc., Scarborough, ME) instrument which uses immersion oil-based optics to image cells smaller than 1 µm in diameter. All samples were imaged using a 50 x 1000 µm flow cell and a flow rate of 0.01 mL/min. Prior to measurements, fresh immersion oil was added to the instrument's optics and the optics were aligned and focused on 1 µm calibration beads. Focusing was performed by priming the system with beads and adjusting the focal plane of the instrument using the micrometer until as many beads were in optimal focus as possible. The flash duration of the instrument was also adjusted to ensure that the average background intensity of the instrument was approximately 160 at the start of each measurement. Samples were imaged by loading 0.1 mL of sample into the instrument's fluidics, priming the flow cell until particles were observed in the system's optics, and analyzing 0.1 mL of sample volume. Image segmentation during these measurements was performed with 15 and 17 dark and light pixel thresholds and rolling

background calibration using the default 10 frames for background estimation. The instrument's fluidics were flushed with ethanol and then water between measurements.

FIM was used to image blood and bacterial cells that could be used to train ConvNet classifiers. To ensure that the cell density was appropriate for the FlowCam Nano, raw blood and bacterial samples were diluted 1:100 with either PBS (blood, *E. coli*, *L. lactis*, and *S. marcescens*) or DMEM (all remaining bacteria) before imaging. Platelet samples obtained from the AcouWash were not diluted beyond the dilution yielded by the platelet separation. For cell types that could be safely handled with biosafety level (BSL) 1 precautions including murine blood, *E. coli*, *S. marcescens,* and *L. lactis*, imaging was performed with the FlowCam Nano on a lab bench. As the remaining organisms required BSL 2 precautions, the remaining organisms were imaged as described but with the FlowCam placed inside a biosafety cabinet. All images were exported from the FlowCam Nano and rescaled to 40 x 40 pixels for further analysis.

8.3.5 Convolutional Neural Networks

ConvNets were trained to distinguish between different types of cell types that may be encountered in a blood sample from a patient with sepsis. Two neural networks were trained for this purpose. The first was trained to distinguish between platelets cells and bacterial cells. The other was instead trained to distinguish between different species of bacteria. Table 7 shows the structure of these ConvNets. Both networks accepted a FIM image and returned the probabilities of the image being each possible class the network was trained on—a blood or bacterial cell for the first network and each of the eight possible species of bacteria for the second. These networks were trained using 16,000 images of platelets and each bacteria type. Training was performed using an Adam optimizer[26] with a 0.001 learning rate and 64-image

Table 7 : ConvNet Structure for blood and bacteria classification
The structure used for both ConvNet classifiers trained in this analysis. The number of features used in the final layers was 2 for the classifier trained to distinguish platelets from bacterial cells and 8 for the classifier used to distinguish different bacteria species

Layer #	Layer Type	# of features	Kernel Size	Activation	Input Shape	Output Shape
1	Convolutional	32	3x3	ReLU	40x40x3	38x38x32
2	Convolutional	32	3x3	ReLU	38x38x32	36x36x32
3	Max Pooling (2x2)	-	-	-	36x36x32	18x18x32
4	Dropout (10% rate)	-	-	-	18x18x32	18x18x32
5	Convolutional	32	3x3	ReLU	18x18x32	16x16x32
6	Convolutional	64	3x3	ReLU	16x16x32	14x14x64
7	Max Pooling (2x2)	-	-	-	14x14x64	7x7x64
8	Dropout (10% rate)	-	-	-	7x7x64	7x7x64
9	Convolutional	64	3x3	ReLU	7x7x64	5x5x64
10	Convolutional	64	3x3	ReLU	5x5x64	3x3x64
11	Flatten	-	-	-	3x3x64	576
12	Dropout (10% rate)	-	-	-	576	576
13	Dense	64	-	ReLU	576	64
14	Dense	2 or 8	-	Softmax	64	2

minibatches. 2,000 images of each cell type were set aside as test images and used to test the performance of the trained classifiers. To test the classifiers, each test image was analyzed with the trained ConvNets to obtain predicted class probabilities that were then used to classify the image. Images were classified as either being a platelet or bacteria using the returned probabilities that the inputted image contained a bacterial cell; if that probability was above a threshold value the cell was classified as a bacteria cell. Three values of this threshold were investigated; 10%, 50%, and 90%. For the bacteria species classifier images were classified as the class for which the network returned the highest probability for.

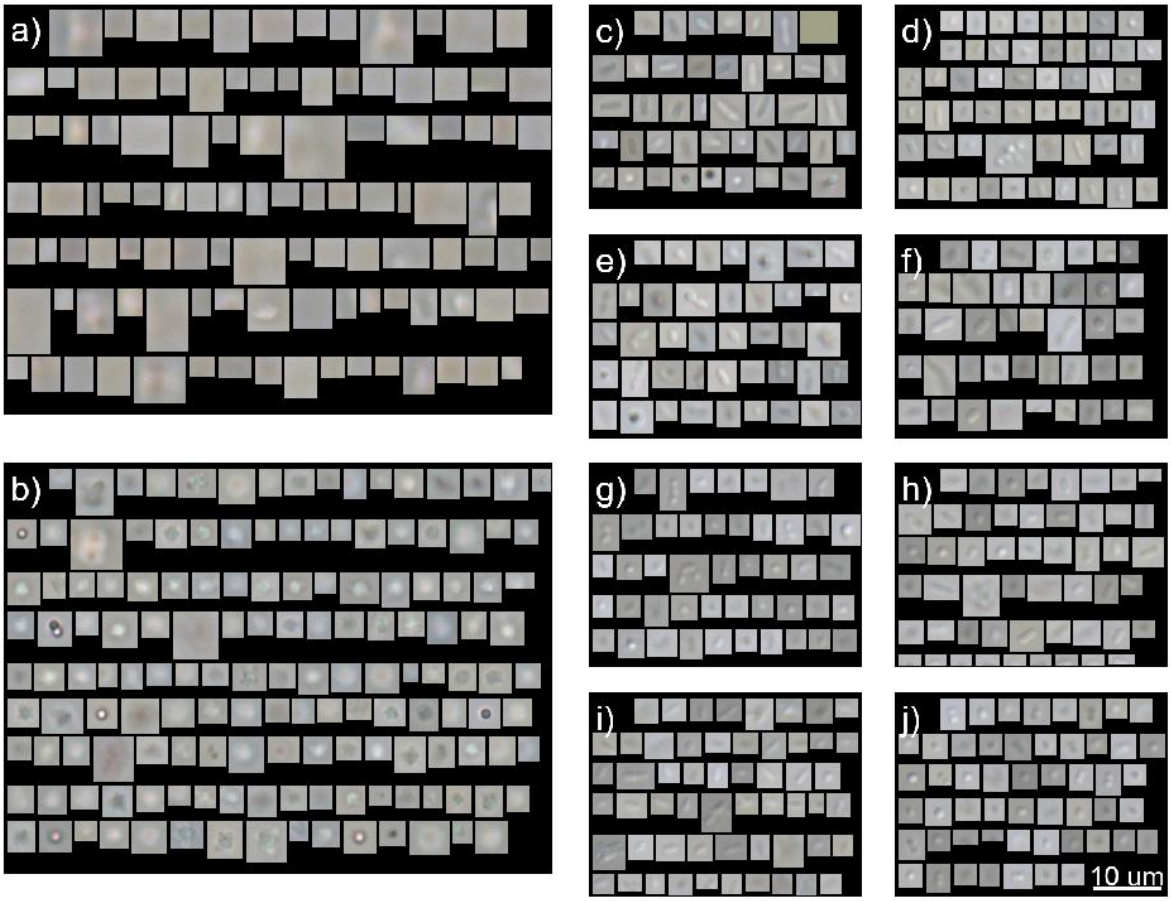

Figure 22: FIM images of blood and bacterial cells
Figure 1: Sample FIM images of (a) whole blood (b) purified platelets (c) *E. coli* (d) *L. lactis* (e) *S. marcescens* (f) *A. baumannii* (g) *E. faecalis* (h) *K. pneumoniae* (i) *P. aeruginosa* (j) *S. aureus.* Images were randomly selected from FIM

8.4 Results

8.4.1 Flow imaging microscopy

Figure 22 shows randomly selected FIM images obtained from whole blood samples, concentrated platelet samples, and cultures of all 8 microorganisms included in this analysis. The large red blood cells present in whole blood sample (Figure 22a) appeared out of focus on the FlowCam when the instrument was focused on 1 µm beads. The size and fuzziness of these

images appeared to interfere with the image segmentation on the FlowCam, resulting in a large number of seemingly-blank images from whole blood samples. These issues were not observed with the smaller platelet and bacterial cells which were more frequently in focus.

As was anticipated, the platelets exhibited similar size, shape, and coloration to the bacteria relative to the large, fuzzy red blood cells (Figure 22b-j). Despite this similarity, the slightly larger, rounded structures of platelets were visually distinguishable from the smaller, elongated structures of the bacteria. Similarly, many of the bacteria species were also visually distinguishable due to the different cell morphologies these species exhibit (e.g. rods for *E. coli* and *P. aeruginosa*, spheres for *E. faecalis* and *S. aureus*).

8.4.2 Convolutional Neural Networks

ConvNets were trained to classify FIM images as either platelets or bacteria. Large red and white blood cells were not used to train this network as it was anticipated that ConvNets could easily identify the dramatic morphology difference between large, out-of-focus red blood cells and bacteria. Table 8 shows a confusion matrix showing the accuracy of the trained classifier on FIM images of platelets and other species of bacteria that were excluded from algorithm training. The classifier yielded great classification accuracy regardless of the threshold, identifying 97-99% of held-out platelet images and ~99% of the bacterial cells of different species. *E. coli* cells yielded the most incorrect classifications with only 96% of the cells being correctly identified as bacteria. The cause of the relatively low classification accuracy on this bacteria species is unknown. As is shown in Table 8, the threshold probability used to classify images can be adjusted to change the relative error rates of this classifier: higher thresholds for

Table 8 : Confusion matrix for blood-bacteria classification

Confusion matrix showing the percent of FIM images from pure cultures of one cell type (rows) that were identified as a platelet or a bacteria cell by the ConvNet (columns). The performance of the classifier at three different thresholds (column headers) is shown here

	10% Threshold		50% Threshold		90% Threshold	
	Platelet	Bacteria	Platelet	Bacteria	Platelet	Bacteria
Platelets	97%	3%	98%	2%	99%	1%
Bacteria (Average)	1%	99%	1%	99%	1%	99%
E. coli	2%	98%	4%	96%	5%	95%
L. lactis	0%	100%	0%	100%	1%	99%
S. marcescens	1%	99%	1%	99%	2%	98%
A. baumannii	1%	99%	1%	99%	1%	99%
E. faecalis	0%	100%	0%	100%	0%	100%
K. pneumoniae	0%	100%	0%	100%	0%	100%
P. aeruginosa	0%	100%	0%	100%	1%	99%
S. aureus	0%	100%	0%	100%	0%	100%

Table 9 : Confusion matrix for bacteria species classification

Confusion matrix showing the percent of FIM images from pure cultures of one bacteria species (rows) that were identified as each of the eight species the ConvNet was trained to detect (columns)

	E. Coli	*L. lactis*	*S. marcescens*	*A. baumannii*	*E. faecalis*	*K. pneumoniae*	*P. aeruginosa*	*S. aureus*
E. coli	99.6%	0.0%	0.2%	0.1%	0.0%	0.0%	0.0%	0.0%
L. lactis	0.3%	91.1%	8.0%	0.1%	0.3%	0.1%	0.2%	0.1%
S. marcescens	1.2%	3.1%	94.9%	0.4%	0.0%	0.0%	0.2%	0.1%
A. baumannii	0.4%	0.0%	0.7%	83.8%	0.7%	9.5%	3.2%	1.7%
E. faecalis	0.0%	0.0%	0.1%	0.4%	58.4%	1.9%	3.6%	35.5%
K. pneumoniae	0.1%	0.0%	0.1%	7.6%	2.3%	74.8%	5.5%	9.5%
P. aeruginosa	0.2%	0.1%	0.3%	4.3%	11.3%	11.7%	65.7%	6.4%
S. aureus	0.1%	0.0%	0.1%	0.5%	4.5%	1.4%	0.9%	92.6%

identifying an FIM image as a bacteria decreases the fraction of misidentified platelets while increasing the fraction of misidentified bacteria and vice versa.

A separate ConvNet was trained to classify images of FIM images identified as bacteria by the previous network. The performance of this network on unseen FIM images of different bacteria species is shown in Table 9. The average accuracy of the classifications yielded by this

ConvNet was 82%. However, we note that the accuracy of the classification varies significantly with organism with species, varying from 58% (*E. faecalis*) to nearly 100% (*E. coli*).

8.5 Discussion

FIM-based methods for diagnosing BSIs potentially offer clinicians a technique to diagnose sepsis rapidly and accurately. To demonstrate proof of concept of this approach, FIM images of platelets and bacteria (including common human pathogens) were collected and used to train ConvNets to distinguish platelets from bacteria and different species of bacteria from each other.

As shown in Table 8, the ConvNet trained to differentiate FIM images of bacteria from those of platelets correctly identified 98% of platelet images and 99% of bacteria images using a 50% threshold. While red and white blood cells were not included in this classifier, it is believed that these large cells would be easy to differentiate from pathogens due to the obvious difference in cell size and morphology when images of these cells are in focus. The relatively high accuracy of this classifier particularly on bacterial cells independent of the threshold suggests that this approach can be used to digitally filter out many of the blood cell images present in a FIM dataset with minimal false negatives on pathogen images.

While this classifier offered high single-image accuracy, a classifier used in practice will need to correctly identify a large fraction of blood cells for this approach to be effective in identifying bacteria. As blood cells greatly outnumber the concentration of pathogen cells in most sepsis cases, a small fraction of incorrectly identified blood cells can easily disguise a few correctly identified pathogenic cells. The classifier here was only able to correctly identify 98% of platelets

at a 50% threshold and, while the fraction of correctly identified platelets can be increased by increasing this threshold, higher overall accuracy on platelet images will likely be required for this approach to be practical. Despite the suboptimal platelet identification, the high accuracy overall accuracy particularly on pathogens suggests that this network may obtain the needed accuracy for BSI diagnosis with additional algorithm refinement. With these improvements, a network similar to the one trained in this analysis could be used to rapidly detect any pathogenic cells that may be in a blood sample—regardless of the species of pathogen or its ability to grow in culture media.

The second ConvNet trained in this analysis was able to correctly predict the species of bacteria in individual unseen FIM images 82% of the time albeit with significant variation in accuracy between organisms. While the accuracy of this classifier could be improved, the relatively high accuracy of the classifier suggests that the ConvNet was generally able to differentiate between FIM images of different bacteria species. We note that the accuracy of the identified causative agent improves if additional pathogenic cells can be collected and assuming a patient is only infected by a single pathogen. As has previously been shown with FIM image classifiers[22] since the ConvNet has over a 50% chance of correctly identifying individual images of each species the accuracy of the ConvNet at predicting the causative agent will increase rapidly as multiple FIM images of the pathogen are collected from a patient. While this simple classifier only includes a small subset of potential pathogens, it should be noted that unlike PCR-based methods FIM and the previous ConvNet can still be used to detect and diagnose infections of bacteria species that this network is not trained against. Regardless, it is still desirable to increase the number of pathogens this classifier can identify in future iterations of this approach.

These results suggest that ConvNets can distinguish between FIM images of different blood and pathogen cells. With the inclusion of additional relevant pathogen species in algorithm training

and improvements in classifier accuracy these networks may be used to identify pathogen cells in an FIM-based BSI diagnosis. Such a diagnosis strategy would provide a fast, sensitive method to diagnose BSIs in patients. FIM can be used to image particles in 0.5 mL of sample per hour and the resulting images can be analyzed by ConvNets nearly instantaneously. It is anticipated that the throughput of FIM will be lower in practice as, due to the high cell concentration, blood samples will need to be diluted to obtain accurate image segmentation during FIM. However, pre-processing whole blood samples using techniques like AcouWash to remove large red and white blood cells from whole blood would minimize the amount of dilution required. Assuming a 1:20 dilution is required to perform this separation, a 50 µL drop of blood can currently be analyzed by a FIM-based BSI diagnosis technique in about two hours of analysis time—much faster than blood culture-based methods and competitive with if not faster than PCR-based methods[15]. While the currently low throughput of this technique does limit the concentration of bacteria that can be detected, the rapid analysis of small blood volumes and potential single-cell sensitivity of ConvNet classifiers is ideal for diagnosing neonatal sepsis. It is also anticipated that the throughput of the technique can be improved with refinements in cell separation and imaging protocols.

While this study focused on training ConvNets to differentiate between blood and pathogen cells, it is anticipated that a similar FIM-based strategy for identifying different cell types may be used for other applications. For example, FIM could be used to diagnose urinary tract infections[27] by using a ConvNet to determine the concentration of different bacterial species like that trained in this study.

8.6 Acknowledgements

We would like to thank Kristen Eller, Natalie Meinerz, Kathryne Walker, Hannah Middlestead, and Alyssa Witeof who assisted with obtaining and analyzing blood samples and bacterial cultures. This work was funded by Yokogawa Fluid Imaging.

8.7 References

1. Rhodes A, Evans LE, Alhazzani W, et al. *Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016*. Vol 45.; 2017. doi:10.1097/CCM.0000000000002255

2. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. *Lancet*. 2018;392(10141):75-87. doi:10.1016/S0140-6736(18)30696-2

3. Gaieski DF, Edwards JM, Kallan MJ, Carr BG. Benchmarking the incidence and mortality of severe sepsis in the united states. *Crit Care Med*. 2013;41(5):1167-1174. doi:10.1097/CCM.0b013e31827c09f8

4. Vincent JL, Marshall JC, Ñamendys-Silva SA, et al. Assessment of the worldwide burden of critical illness: The Intensive Care Over Nations (ICON) audit. *Lancet Respir Med*. 2014;2(5):380-386. doi:10.1016/S2213-2600(14)70061-X

5. Wynn JL. Defining neonatal sepsis. *Curr Opin paediatr*. 2016;28(2):135-140. doi:10.1097/MOP.0000000000000315.Defining

6. Barton L, Hodgman JE, Pavlova Z. Causes of death in the extremely low birth weight infant. *Pediatrics*. 1999;103(2):446-451. doi:10.1542/peds.103.2.446

7. Seymour CW, Gesten F, Prescott HC, et al. Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. *N Engl J Med*. 2017;376(23):2235-2244. doi:10.1056/nejmoa1703058

8. Wain J, Diep TS, Ho VA, et al. Quantitation of Bacteria in Blood of Typhoid Fever Patients and Relationship between Counts and Clinical Features , Transmissibility , and Antibiotic Resistance Quantitation of Bacteria in Blood of Typhoid Fever Patients and Relationship between Counts an. 1998;36(6):1683-1687.

9. Opota O, Jaton K, Greub G. Microbial diagnosis of bloodstream infection: Towards molecular diagnosis directly from blood. *Clin Microbiol Infect*. 2015;21(4):323-331. doi:10.1016/j.cmi.2015.02.005

10. Thorpe TC, Wilson ML, Turner JE, et al. BacT/Alert: An automated colorimetric microbial detection system. *J Clin Microbiol*. 1990;28(7):1608-1612.

11. Totty H, Ullery M, Spontak J, et al. A controlled comparison of the BacT/ALERT® 3D and VIRTUO™ microbial detection systems. *Eur J Clin Microbiol Infect Dis*. 2017;36(10):1795-1800. doi:10.1007/s10096-017-2994-8

12. Biondi EA, Mischler M, Jerardi KE, et al. Blood culture time to positivity in febrile infants with bacteremia. *JAMA Pediatr*. 2014;168(9):844-849. doi:10.1001/jamapediatrics.2014.895

13. Westh H, Lisby G, Breysse F, et al. Multiplex real-time PCR and blood culture for identification of bloodstream pathogens in patients with suspected sepsis. *Clin Microbiol Infect*. 2009;15(6):544-551. doi:10.1111/j.1469-0691.2009.02736.x

14. Lamoth F, Jaton K, Prod'hom G, et al. Multiplex blood PCR in combination with blood cultures for improvement of microbiological documentation of infection in febrile neutropenia. *J Clin Microbiol*. 2010;48(10):3510-3516. doi:10.1128/JCM.00147-10

15. Sinha M, Jupe J, Mack H, Coleman TP, Lawrence SM, Fraley I. Emerging Technologies for Molecular Diagnosis of Sepsis. *Clin Microbiol Rev*. 2018;31(2):1-26.

16. Sieracki CK, Sieracki ME, Yentsch CS. An imaging-in-flow system for automated analysis of marine microplankton. *Mar Ecol Prog Ser*. 1998;168:285-296. doi:10.3354/meps168285

17. Sharma DK, King D, Oma P, Merchant C. Micro-Flow Imaging: Flow Microscopy Applied to Sub-visible Particulate Analysis in Protein Formulations. *AAPS J*. 2010;12(3):455-464. doi:10.1208/s12248-010-9205-1

18. Sediq AS, Klem R, Nejadnik MR, Meij P, Jiskoot W. Label-Free, Flow-Imaging Methods for Determination of Cell Concentration and Viability. *Pharm Res*. 2018;35(8). doi:10.1007/s11095-018-2422-5

19. Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions. *IEEE Conf Comput Vis Pattern Recognit*. 2015. doi:10.1109/CVPR.2015.7298594

20. Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2015;07-12-June:815-823. doi:10.1109/CVPR.2015.7298682

21. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118. doi:10.1038/nature21056

22. Calderon CP, Daniels AL, Randolph TW. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. *J Pharm Sci*. 2018;107(4):999-1008. doi:10.1016/j.xphs.2017.12.008

23. Daniels AL, Calderon CP, Randolph TW. Machine Learning & Statistical Analyses for Extracting and Characterizing “Fingerprints” of Antibody Aggregation at Container Interfaces from Flow Microscopy Images. *Biotechnol Bioeng*. 2020.

24. Pitt W, Alizadeh M, Husseini GA, et al. Rapid Separation of Bacteria from Blood—Review and Outlook. 2016;32(4):823-839. doi:10.5588/ijtld.16.0716.Isoniazid

25. Ohlsson P, Petersson K, Augustsson P, Laurell T. Acoustic impedance matched buffers enable separation of bacteria from blood cells at high cell concentrations. *Sci Rep*. 2018;8(1):1-11. doi:10.1038/s41598-018-25551-0

26. Kingma DP, Ba JL. Adam: A method for stochastic optimization. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. 2015:1-15.

27. Chu CM, Lowder JL. Diagnosis and treatment of urinary tract infections across age groups. *Am J Obstet Gynecol*. 2018;219(1):40-51. doi:10.1016/j.ajog.2017.12.231

Chapter 9: Conclusions

9.1 Overview

The previous chapters describe approaches that rely on flow imaging microscopy (FIM) and various machine learning methods to analyze the morphology of particles present in a liquid sample. The resulting algorithms were then used to analyze FIM images of protein aggregates formed under different conditions such as different accelerated stability stresses or primary containers. These approaches were also adapted to analyze imaging data from a different modality (imaging flow cytometry) and to analyze cells rather than protein aggregates. While these studies investigated diverse applications of FIM and machine learning algorithms, there were a few overarching conclusions from these studies which will be described in this chapter.

9.2 FIM images contain particle morphology information that is characteristic of different particle types

Each of the algorithms described in chapter 2 is trained to distinguish between FIM images of particles from samples that were treated in different ways (e.g. protein formulations exposed to different stresses, cultures of different organisms). It was hypothesized that the particles in

these samples would exhibit distinguishable particle morphologies, resulting in FIM images that could be differentiated between samples using machine learning algorithms.

The performance of the developed algorithms was consistent with this hypothesis. The machine learning algorithms developed as part of this thesis were able to differentiate between particles from different samples. This suggests that information recorded in raw FIM images could be used by these algorithms to differentiate between particle types. Furthermore, the results in Chapters 5-6 also suggest that these images are differentiable at least in part due to underlying differences in particle morphology. When analyzing FIM images using the convolutional neural network (ConvNet)-based particle comparison algorithm, the positions of raw FIM images in the resulting ConvNet-derived embedding space were often found to correlate with human-observable particle morphology features. While it is possible that other uninterpretable image features were also used by ConvNet-based algorithms to compute these image embeddings[1], these results suggest that ConvNet-based algorithms likely used the underlying particle morphology information captured in these images in part to help differentiate between particle types. The sensitivity of these algorithms to changes in particle morphology captured by FIM images allowed us to use these algorithms to analyze FIM images to differentiate between protein aggregates and cells based on their morphology.

9.3 Subvisible protein aggregates formed by different stress conditions exhibit different morphologies

Several of the studies presented in this thesis (see Chapters 2, 4-6) used FIM and machine learning methods to compare protein aggregates formed under different accelerated stress conditions such as freeze-thaw, shaking, and tumbling. These different stress conditions

frequently yielded dramatic differences in particle morphology; particles generated by these different stresses were often distinguishable by eye (e.g. Figure 11) and using machine learning approaches (e.g. Figure 12). This result was expected; as these different stresses likely induce aggregation through dramatically different mechanisms, the resulting aggregates assembled through these mechanisms were likely to exhibit drastically different particle morphologies characteristic of these different mechanisms. It was also observed that subtle changes in formulation conditions (e.g. container type in Chapter 4, secondary packaging in Chapter 5) can also subtly influence these aggregation mechanisms, generating slightly different particle populations than those present without these changes.

Because of this effect, particle morphologies could be used as a tool to analyze the dominant mechanisms of aggregation in samples. This approach was used to identify the dominant mechanism of aggregation during shipping stress and to determine if fibrillation occurred during cavitation. Comparing aggregates formed by shipping to those formed by low and high g-force stresses samples experience during shipping revealed that low g-force stresses are the primary cause of aggregation from shipping. A similar analysis revealed that insulin and amyloid beta formulations form a similar aggregate population after exposure to different cavitation-inducing stress and elevated temperature incubation, suggesting that cavitation may generate amyloid fibrils of these proteins.

These results suggest that these algorithms may be useful in diagnosing process upsets in manufacturing processes for these protein formulations. As aggregate morphologies are sensitive to changes in their mechanisms of formation, deviations in manufacturing conditions that trigger aggregation will result in a new aggregate morphology characteristic of that upset. The particle comparison algorithm described in Chapters 2 and 4 was designed to detect these particle morphology changes. This algorithm can be trained on particles normally generated by

the manufacturing process and used to compare particles in new batches of product against these particles. These comparisons can be used to detect changes in particle morphology in these new batches that indicate a change in the underlying mechanisms of aggregation (i.e. a process upset). If a deviation in particle morphology is detected, either of the ConvNet methods described here could then be used to compare the new particle population to those made by known stresses or process upsets to help identify the root cause of the upset.

9.4 Machine learning approaches developed for FIM images can be applied to similar imaging modalities

While most of the analysis performed here used FIM to obtain particle images, the algorithms described here are agnostic to the imaging platform and may be applied to particle images from other imaging modalities. Chapter 7 described how these algorithms can be adapted for data returned by imaging flow cytometry (IFC), a similar imaging platform to FIM that returns both light and fluorescence microscopy images for each particle. The results obtained from this study not only suggested that these algorithms could be adapted to IFC images but that the additional particle morphology available from the fluorescence microscopy images resulted in better differentiation between protein aggregates generated by different stresses. While the image-centric data returned by IFC made it simple to adapt ConvNet-based algorithms to analyze the data it collects, it is anticipated that these machine learning approaches can also be tailored to other imaging platforms like background membrane imaging[2]. Doing so may allow these algorithms to include additional particle morphology information in this analysis that is not available from FIM alone, potentially enabling the algorithm to better distinguish between particle types.

9.5 FIM and machine learning approaches can be used to differentiate cells of different species

Although the algorithms described here were developed for analyzing protein formulations, these methods were also successfully used to differentiate between cells of different organisms. Similar ConvNet-based methods used to classify protein aggregates from different stresses described in Chapter 2 were also successfully used to differentiate between platelets and different species of bacteria as described in Chapter 8. The ability for these algorithms to differentiate between cell species may be useful for several different applications. As described in chapter 8, FIM and ConvNets may be used to identify pathogen cells in a blood sample as a rapid, sensitive sepsis diagnosis strategy. A similar approach could also be used to diagnose urinary tract infections by detecting pathogens in a patient's urine in a similar but much higher throughput fashion than currently-used microscopy techniques[3]. FIM may also be useful for monitoring cells found in non-medical applications such as monitoring cell-based medical product production[4] or detecting microbes in water samples[5].

9.6 References

1. Wang H, Wu X, Huang Z, Xing EP. High Frequency Component Helps Explain the Generalization of Convolutional Neural Networks. 2019;(Remark 1):8684-8694. http://arxiv.org/abs/1905.13545.

2. Helbig C, Ammann G, Menzen T, Friess W, Wuchner K, Hawe A. Backgrounded Membrane Imaging (BMI) for High-Throughput Characterization of Subvisible Particles During Biopharmaceutical Drug Product Development. *J Pharm Sci*. 2019:1-13. doi:10.1016/j.xphs.2019.03.024

3. Chu CM, Lowder JL. Diagnosis and treatment of urinary tract infections across age groups. *Am J Obstet Gynecol*. 2018;219(1):40-51. doi:10.1016/j.ajog.2017.12.231

4. Sediq AS, Klem R, Nejadnik MR, Meij P, Jiskoot W. Label-Free, Flow-Imaging Methods for

Determination of Cell Concentration and Viability. *Pharm Res*. 2018;35(8). doi:10.1007/s11095-018-2422-5

5. Sieracki CK, Sieracki ME, Yentsch CS. An imaging-in-flow system for automated analysis of marine microplankton. *Mar Ecol Prog Ser*. 1998;168:285-296. doi:10.3354/meps168285

Bibliography

Ahmadi, Maryam, Christine J. Bryson, Edward A. Cloake, Katie Welch, Vasco Filipe, Stefan Romeijn, Andrea Hawe, Wim Jiskoot, Matthew P. Baker, and Mark H. Fogg. 2015. “Small Amounts of Sub-Visible Aggregates Enhance the Immunogenic Potential of Monoclonal Antibody Therapeutics.” *Pharmaceutical Research* 32 (4): 1383–94. https://doi.org/10.1007/s11095-014-1541-x.

Anderson, T. W., and D. A. Darling. 1954. “A Test of Goodness of Fit.” *Journal of the American Statistical Association* 49 (268): 765–69. https://doi.org/10.1017/CBO9781107415324.004.

Arakawa, Tsutomu, Daisuke Ejima, and Teruo Akuta. 2017. “Protein Aggregation under High Concentration/Density State during Chromatographic and Ultrafiltration Processes.” *International Journal of Biological Macromolecules* 95: 1153–58. https://doi.org/10.1016/j.ijbiomac.2016.11.005.

Arindra, Arnaud, Adiyoso Setio, Francesco Ciompi, Geert Litjens, Paul Gerke, Colin Jacobs, Sarah J Van Riel, et al. 2016. “Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks.” *Ieee Transactions on Medical Imaging* 35 (5): 1160–69. https://doi.org/10.1109/TMI.2016.2536809.

Arsiccio, Andrea, and Roberto Pisano. 2017. “Stability of Proteins in Carbohydrates and Other Additives during Freezing: The Human Growth Hormone as a Case Study.” *Journal of Physical Chemistry B* 121: 8652–60.

Barandun, S., F. Jeunet, P. Kistler, and H. Isliker. 1962. “Intravenous Administration of Human Gamma-Globulin.” *Vox Sanguinis* 7 (2): 157–74.

Barnard, James G., David Kahn, David Cetlin, Theodore W. Randolph, and John F. Carpenter. 2014. “Investigations into the Fouling Mechanism of Parvovirus Filters during Filtration of Freeze-Thawed MAb Drug Substance Solutions.” *Journal of Pharmaceutical Sciences*. https://doi.org/10.1002/jps.23881.

Barnard, James G., Satish Singh, Theodore W. Randolph, and John F. Carpenter. 2011. “Subvisible Particle Counting Provides a Sensitive Method of Detecting and Quantifying Aggregation of Monoclonal Antibody Caused by Freeze-Thawing: Insights into the Roles of Particles in the Protein Aggregation Pathway.” *J Pharm Sci 2011* 100 (2): 492–503.

Barton, Lorayne, Joan E. Hodgman, and Zdena Pavlova. 1999. “Causes of Death in the Extremely Low Birth Weight Infant.” *Pediatrics* 103 (2): 446–51. https://doi.org/10.1542/peds.103.2.446.

Bee, Jared S., Theodore W. Randolph, John F. Carpenter, Steven M. Bishop, and Mariana N. Dimitrova. 2011. “Effects of Surfaces and Leachables on the Stability of Biopharmaceuticals.” *Journal of Pharmaceutical Sciences* 100 (10): 4158–70. https://doi.org/10.1002/jps.22597.

Bee, Jared S., Daniel K. Schwartz, Siwar Trabelsi, Erwin Freund, Jennifer L. Stevenson, John F. Carpenter, and Theodore W. Randolph. 2012. “Production of Particles of Therapeutic Proteins at the Air-Water Interface during Compression/Dilation Cycles.” *Soft Matter* 8 (40): 10329–35. https://doi.org/10.1039/c2sm26184g.

Beers, Miranda M.C. Van, and Muriel Bardor. 2012. “Minimizing Immunogenicity of Biopharmaceuticals by Controlling Critical Quality Attributes of Proteins.” *Biotechnology Journal* 7 (12): 1473–84. https://doi.org/10.1002/biot.201200065.

Beers, Miranda M.C. Van, Francesca Gilli, Huub Schellekens, Theodore W. Randolph, and Wim Jiskoot. 2012. “Immunogenicity of Recombinant Human Interferon Beta Interacting with Particles of Glass, Metal, and Polystyrene.” *Journal of Pharmaceutical Sciences* 101 (1): 187–99. https://doi.org/10.1002/jps.22744.

Beers, Miranda M C van, Wim Jiskoot, and Huub Schellekens. 2010. “On the Role of Aggregates in the Immunogenicity of Recombinant Human Interferon Beta in Patients with Multiple Sclerosis.” *Journal of Interferon & Cytokine Research* 30 (10): 767–75.

Bessa, Juliana, Sabine Boeckle, Hermann Beck, Thomas Buckel, Sonja Schlicht, Martin Ebeling, Anna Kiialainen, et al. 2015. “The Immunogenicity of Antibody Aggregates in a Novel Transgenic Mouse Model.” *Pharmaceutical Research* 32 (7): 2344–59. https://doi.org/10.1007/s11095-015-1627-0.

Bhatnagar, Bakul S., Robin H. Bogner, and Michael J. Pikal. 2007. “Protein Stability during Freezing: Separation of Stresses and Mechanisms of Protein Stabilization.” *Pharmaceutical Development and Technology* 12 (5): 505–23. https://doi.org/10.1080/10837450701481157.

Biondi, Eric A., Matthew Mischler, Karen E. Jerardi, Angela M. Statile, Jason French, Rianna Evans, Vivian Lee, et al. 2014. “Blood Culture Time to Positivity in Febrile Infants with Bacteremia.” *JAMA Pediatrics* 168 (9): 844–49. https://doi.org/10.1001/jamapediatrics.2014.895.

Bojarski, Mariusz, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Prasoon Goyal, Lawrence D. Jackel, et al. 2016. “End to End Learning for Self-Driving Cars,” 1–9. http://arxiv.org/abs/1604.07316.

Buskey, Edward J., and Cammie J. Hyatt. 2006. “Use of the FlowCAM for Semi-Automated Recognition and Enumeration of Red Tide Cells (Karenia Brevis) in Natural Plankton Samples.” *Harmful Algae* 5 (6): 685–92. https://doi.org/10.1016/j.hal.2006.02.003.

Calderon, Christopher P., Austin L. Daniels, and Theodore W. Randolph. 2018. “Deep

Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations." *Journal of Pharmaceutical Sciences* 107 (4): 999–1008. https://doi.org/10.1016/j.xphs.2017.12.008.

Cecconi, Maurizio, Laura Evans, Mitchell Levy, and Andrew Rhodes. 2018. "Sepsis and Septic Shock." *The Lancet* 392 (10141): 75–87. https://doi.org/10.1016/S0140-6736(18)30696-2.

Chi, Eva Y. 2004. "Protein Aggregation in Aqueous Solution -- Mechanism, Thermodynamics, and Kinetics."

Chi, Eva Y., Sampathkumar Krishnan, Theodore W. Randolph, and John F. Carpenter. 2003. "Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation." *Pharmaceutical Research* 20 (9): 1325–36. https://doi.org/10.1023/A:1025771421906.

Chisholm, Carly Fleagle, Abby E. Baker, Kaitlin R. Soucie, Raul M. Torres, John F. Carpenter, and Theodore W. Randolph. 2016. "Silicone Oil Microdroplets Can Induce Antibody Responses Against Recombinant Murine Growth Hormone in Mice." *Journal of Pharmaceutical Sciences* 105 (5): 1623–32. https://doi.org/10.1016/j.xphs.2016.02.019.

Chisholm, Carly Fleagle, Kaitlin R. Soucie, Jane S. Song, Pamela Strauch, Raul M. Torres, John F. Carpenter, Jack A. Ragheb, and Theodore W. Randolph. 2017. "Immunogenicity of Structurally Perturbed Hen Egg Lysozyme Adsorbed to Silicone Oil Microdroplets in Wild-Type and Transgenic Mouse Models." *Journal of Pharmaceutical Sciences* 106 (6): 1519–27.

Chiti, Fabrizio, and Christopher M. Dobson. 2006. "Protein Misfolding, Functional Amyloid, and Human Disease." *Annual Review of Biochemistry* 75: 333–66. https://doi.org/10.1146/annurev.biochem.75.101304.123901.

Chu, Christine M., and Jerry L. Lowder. 2018. "Diagnosis and Treatment of Urinary Tract Infections across Age Groups." *American Journal of Obstetrics and Gynecology* 219 (1): 40–51. https://doi.org/10.1016/j.ajog.2017.12.231.

Cohen, Samuel I. A., Sara Linse, Leila M. Luheshi, Erik Hellstrand, Duncan A. White, Luke Rajah, Daniel E. Otzen, Michele Vendruscolo, Christopher M. Dobson, and Tuomas P. J. Knowles. 2013. "Proliferation of Amyloid-B42 Aggregates Occurs through a Secondary Nucleation Mechanism." *Proceedings of the National Academy of Sciences of the United States of America* 110 (24): 9758–63. https://doi.org/10.1073/pnas.1218402110.

Cordes, Amanda A., John F. Carpenter, and Theodore W. Randolph. 2012. "Accelerated Stability Studies of Abatacept Formulations: Comparison of Freeze–Thawing- and Agitation-Induced Stresses." *Journal of Pharmaceutical Sciences* 101 (7): 2307–15.

Corvari, Vincent, Linda O. Narhi, Thomas M. Spitznagel, Nataliya Afonina, Shawn Cao, Patricia Cash, Irene Cecchini, et al. 2015. "Subvisible (2-100 Mm) Particle Analysis during Biotherapeutic Drug Product Development: Part 2, Experience with the Application of Subvisible Particle Analysis." *Biologicals* 43 (6): 457–73. https://doi.org/10.1016/j.biologicals.2015.07.011.

Daniels, Austin L., Christopher P. Calderon, and Theodore W. Randolph. 2020. "Machine

Learning & Statistical Analyses for Extracting and Characterizing 'Fingerprints' of Antibody Aggregation at Container Interfaces from Flow Microscopy Images." *Biotechnology and Bioengineering*.

Daniels, Austin L., and Theodore W. Randolph. 2018. "Flow Microscopy Imaging Is Sensitive to Characteristics of Subvisible Particles in Peginesatide Formulations Associated With Severe Adverse Reactions." *Journal of Pharmaceutical Sciences* 107 (5): 1313–21. https://doi.org/10.1016/j.xphs.2018.01.015.

Darling, D. A. 1957. "The Kolmogorov-Smirnov , Cramer-von Mises Tests." *The Annals of Mathematical Statistics* 28 (4): 823–38.

Dempster, AP, and NM Laird. 1977. *Maximum Likelihood from Incomplete Data via the EM Algorithm*. *Journal of the Royal Statistical Society. Series B (* …. Vol. 39. http://www.ams.org/leavingmsn?url=http://links.jstor.org/sici?sici=0035-9246(1977)39:1%3C1:MLFIDV%3E2.0.CO;2-Z&origin=MSN.

Dill, Sarah, Kyle Brees, Andrew Stahly, Elizabeth Cheng, John Carpenter, and Liron Caplan. 2020. "Mechanical Shock During Shipping of Medications: The Roles of Packaging and Transportation Vendors." *Journal of Pharmaceutical Sciences* 109 (1): 670–76. https://doi.org/10.1016/j.xphs.2019.10.050.

Dong, Jinping, Allison Hubel, John C. Bischof, and Alptekin Aksan. 2009. "Freezing-Induced Phase Separation and Spatial Microheterogeneity in Protein Solutions." *Journal of Physical Chemistry B* 113 (30): 10081–87. https://doi.org/10.1021/jp809710d.

Dugas, Charles, Yoshua Bengio, François Bélisle, Claude Nadeau, and René Garcia. 2001. "Incorporating Second-Order Functional Knowledge for Better Option Pricing." *Advances in Neural Information Processing Systems*.

Ellis, Elliot F., and Christopher S. Henney. 1969. "Adverse Reactions Following Administration of Human Gamma Globulin." *Journal of Allergy* 43 (1): 45–54. https://doi.org/10.1016/0021-8707(69)90019-7.

Engelsman, John Den, Patrick Garidel, Ronald Smulders, Hans Koll, Bryan Smith, Stefan Bassarab, Andreas Seidl, Otmar Hainzl, and Wim Jiskoot. 2011. "Strategies for the Assessment of Protein Aggregates in Pharmaceutical Biotech Product Development." *Pharmaceutical Research* 28 (4): 920–33. https://doi.org/10.1007/s11095-010-0297-1.

Ennis, R. D., R. Pritchard, C. Nakamura, M. Coulon, T. Yang, G. C. Visor, and W. A. Lee. 2001. "Glass Vials for Small Volume Parenterals: Influence of Drug and Manufacturing Processes on Glass Delamination." *Pharmaceutical Development and Technology* 6 (3): 393–405. https://doi.org/10.1081/PDT-100002248.

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." *Nature* 542 (7639): 115–18. https://doi.org/10.1038/nature21056.

FDA. 2012. "Highlights Of Perscribing Information: Omontys."

Flannigan, David J., and Kenneth S. Suslick. 2005. "Plasma Formation and Temperature

Measurement during Single-Bubble Cavitation." *Nature* 434 (7029): 52–55. https://doi.org/10.1038/nature03361.

Flint, Edward B., and Kenneth S. Suslick. 1991. "The Temperature of Cavitation." *Science* 253 (5026): 1397–99. https://doi.org/10.1126/science.253.5026.1397.

Fradkin, Amber Haynes, John F. Carpenter, and Theodore W. Randolph. 2009. "Immunogenicity of Aggregates of Recombinant Human Growth Hormone in Mouse Models." *Journal of Pharmaceutical Sciences* 98 (9): 3247–64. https://doi.org/10.1002/jps.21834.

———. 2011. "Glass Particles as an Adjuvant: A Model for Adverse Immunogenicity of Therapeutic Proteins." *Journal of Pharmaceutical Sciences* 100 (11): 4953–64.

Franc, Jean-Pierre, and Jean-Marie Michel. 2006. *Fundamentals of Cavitation*. Vol. 76. Springer science & Business media.

Freitag, Angelika J., Maliheh Shomali, Stylianos Michalakis, Martin Biel, Michael Siedler, Zehra Kaymakcalan, John F. Carpenter, Theodore W. Randolph, Gerhard Winter, and Julia Engert. 2015. "Investigation of the Immunogenicity of Different Types of Aggregates of a Murine Monoclonal Antibody in Mice." *Pharmaceutical Research* 32 (2): 430–44. https://doi.org/10.1007/s11095-014-1472-6.

Gaieski, David F., J. Matthew Edwards, Michael J. Kallan, and Brendan G. Carr. 2013. "Benchmarking the Incidence and Mortality of Severe Sepsis in the United States." *Critical Care Medicine* 41 (5): 1167–74. https://doi.org/10.1097/CCM.0b013e31827c09f8.

Gambe-Gilbuena, Arni, Yuriko Shibano, Elena Krayukhina, Tetsuo Torisu, and Susumu Uchiyama. 2020. "Automatic Identification of the Stress Sources of Protein Aggregates Using Flow Imaging Microscopy Images." *Journal of Pharmaceutical Sciences* 109 (1): 614–23. https://doi.org/10.1016/j.xphs.2019.10.034.

Gerhardt, Alana, Nicole R. McGraw, Daniel K. Schwartz, Jared S. Bee, John F. Carpenter, and Theodore W. Randolph. 2014. "Protein Aggregation and Particle Formation in Prefilled Glass Syringes." *Journal of Pharmaceutical Sciences* 103 (6): 1601–12. https://doi.org/10.1002/jps.23973.

Gikanga, Bension, Ada Hui, and Yuh-Fun Maa. 2017. "Mechanistic Investigation on Grinding-Induced Subvisible Particle Formation during Mixing and Filling of Monoclonal Antibody Formulations." https://doi.org/10.5731/pdajpst.2017.007732.

Godin, D. S., E. M. Frohman, B. Hurwitz, P. W. O'Connor, J. J. Oger, A. T. Reder, and J. C. Stevens. 2007. "Neutralizing Antibodies to Interferon Beta: Assesment of Their Clinical and Radiographic Impact: An Evidence Report - Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology." *Neurology* 68 (13): 977–84.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. *MIT Press*. http://www.deeplearningbook.org/.

Ha, Emily, Wei Wang, and Y. John Wang. 2002. "Peroxide Formation in Polysorbate 80 and

Protein Stability." *Journal Of Pharmaceutical Sciences* 91: 2252–64.

Hartung, Hans Peter, F. Munschauer, and H. Schellekens. 2005. "Significance of Neutralizing Antibodies to Interferon Beta during Treatment of Multiple Sclerosis: Expert Opinions Based on the Proceedings of an International Consensus Conference." *European Journal of Neurology* 12 (8): 588–601. https://doi.org/10.1111/j.1468-1331.2005.01104.x.

Hawe, Andrea, Marc Sutter, and Wim Jiskoot. 2008. "Extrinsic Fluorescent Dyes as Tools for Protein Characterization." *Pharmaceutical Research* 25 (7): 1487–99. https://doi.org/10.1007/s11095-007-9516-9.

Helbig, Constanze, Gregor Ammann, Tim Menzen, Wolfgang Friess, Klaus Wuchner, and Andrea Hawe. 2019. "Backgrounded Membrane Imaging (BMI) for High-Throughput Characterization of Subvisible Particles During Biopharmaceutical Drug Product Development." *Journal of Pharmaceutical Sciences*, 1–13. https://doi.org/10.1016/j.xphs.2019.03.024.

Her, Cheng, and John F. Carpenter. 2020. "Effects of Tubing Type, Formulation, and Postpumping Agitation on Nanoparticle and Microparticle Formation in Intravenous Immunoglobulin Solutions Processed With a Peristaltic Filling Pump." *Journal of Pharmaceutical Sciences* 109 (1): 739–49. https://doi.org/10.1016/j.xphs.2019.05.013.

Hermans, Alexander, Lucas Beyer, and Bastian Leibe. 2017. "In Defense of the Triplet Loss for Person Re-Identification." http://arxiv.org/abs/1703.07737.

Hermeling, Suzanne, Daan J A Crommelin, Huub Schellekens, and Wim Jiskoot. 2004. "Structure-Immunogenicity Relationships of Therapeutic Proteins." *Pharmaceutical Research* 21 (6): 897–903. https://doi.org/10.1023/B:PHAM.0000029275.41323.a6.

Hochuli, E. 1997. "Interferon Immunogenicity: Technical Evaluation of Interferon Alpha 2a." *Journal of Interferon & Cytokine Research* 17: S15–21.

Hong, Yongmiao, and Haitao Li. 2005. "Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates." *Review of Financial Studies* 18 (1): 37–84. https://doi.org/10.1093/rfs/hhh006.

Ioffe, Sergey, and Christian Szegedy. 2015. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." *Proceedings of the 32nd International Conference on Machine Learning*. https://doi.org/10.1007/s13398-014-0173-7.2.

Iwaya, Keiichi, Tamotsu Zako, Junta Fukunaga, Karin Margareta Sörgjerd, Kentaro Ogata, Koichiro Kogure, Hiroshi Kosano, et al. 2019. "Toxicity of Insulin-Derived Amyloidosis : A Case Report," 1–6.

Jiskoot, Wim, Grzegorz Kijanka, Theodore W. Randolph, John F. Carpenter, Atanas V. Koulov, Hanns Christian Mahler, Marisa K. Joubert, Vibha Jawa, and Linda O. Narhi. 2016. "Mouse Models for Assessing Protein Immunogenicity: Lessons and Challenges." *Journal of Pharmaceutical Sciences* 105 (5): 1567–75. https://doi.org/10.1016/j.xphs.2016.02.031.

Joubert, Marisa K., Martha Hokom, Catherine Eakin, Lei Zhou, Meghana Deshpande, Matthew P. Baker, Theresa J. Goletz, et al. 2012. "Highly Aggregated Antibody Therapeutics Can

Enhance the in Vitro Innate and Late-Stage T-Cell Immune Responses." *Journal of Biological Chemistry* 287 (30): 25266–79. https://doi.org/10.1074/jbc.M111.330902.

Joubert, Marisa K., Quanzhou Luo, Yasser Nashed-Samuel, Jette Wypych, and Linda O. Narhi. 2011. "Classification and Characterization of Therapeutic Antibody Aggregates." *Journal of Biological Chemistry* 286 (28): 25118–33. https://doi.org/10.1074/jbc.M110.160457.

Justel, Ana, Daniel Peña, and Rubén Zamar. 1997. "A Multivariate Kolmogorov-Smirnov Test of Goodness of Fit." *Statistics & Probability Letters* 35 (3): 251–59. https://doi.org/10.1016/S0167-7152(97)00020-5.

Kalonia, Cavan, Ozan S. Kumru, Indira Prajapati, Roman Mathaes, Julia Engert, Shuxia Zhou, C. Russell Middaugh, and David B. Volkin. 2015. "Calculating the Mass of Subvisible Protein Particles with Improved Accuracy Using Microflow Imaging Data." *Journal of Pharmaceutical Sciences* 104 (2): 536–47. https://doi.org/10.1002/jps.24156.

Kasimbeg, Priya N.O., Fook Chiong Cheong, David B. Ruffner, Jaroslaw M. Blusewicz, and Laura A. Philips. 2018. "Holographic Characterization of Protein Aggregates in the Presence of Silicone Oil and Surfactants." *Journal of Pharmaceutical Sciences* 108 (1): 155–61. https://doi.org/10.1016/j.xphs.2018.10.002.

Kerwin, Bruce A. 2008. "Polysorbates 20 and 80 Used in the Formulation of Protein Biotherapeutics: Structure and Degradation Pathways." *Journal of Pharmaceutical Sciences* 97 (8): 2924–35.

Kiese, Sylvia, Astrid PappenBerger, Wolfgamg Friess, and Hanns-Christian Mahler. 2008. "Shaken, Not Stirred: Mechanical Stress Testing of an lgG1 Antibody." *Journal of Pharmaceutical Sciences* 97 (10): 4347–66. https://doi.org/10.1002/jps.

Kingma, Diederik P., and Jimmy Lei Ba. 2015. "Adam: A Method for Stochastic Optimization." *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.

Kingma, Diederik P., and Max Welling. 2014. "Auto-Encoding Variational Bayes." 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings, no. Ml: 1–14.

Kobayashi, Hirofumi, Cheng Lei, Yi Wu, Chun Jung Huang, Atsushi Yasumoto, Masahiro Jona, Wenxuan Li, et al. 2019. "Intelligent Whole-Blood Imaging Flow Cytometry for Simple, Rapid, and Cost-Effective Drug-Susceptibility Testing of Leukemia." *Lab on a Chip* 19 (16): 2688–98. https://doi.org/10.1039/c8lc01370e.

Kolhe, Parag, Elizabeth Amend, and Satish K. Singh. 2010. "Impact of Freezing on PH of Buffered Solutions and Consequences for Monoclonal Antibody Aggregation." *Biotechnology Progress* 26 (3): 727–33. https://doi.org/10.1002/btpr.377.

Kotarek, Joseph, Christine Stuart, Silvia H. De Paoli, Jan Simak, Tsai Lien Lin, Yamei Gao, Mikhail Ovanesov, et al. 2016. "Subvisible Particle Content, Formulation, and Dose of an Erythropoietin Peptide Mimetic Product Are Associated with Severe Adverse Postmarketing Events." *Journal of Pharmaceutical Sciences* 105 (3): 1023–27. https://doi.org/10.1016/S0022-3549(15)00180-X.

Krayukhina, Elena, Kouhei Tsumoto, Susumu Uchiyama, and Kiichi Fukui. 2015. “Effects of Syringe Material and Silicone Oil Lubrication on the Stability of Pharmaceutical Proteins.” *Journal of Pharmaceutical Sciences* 104 (2): 527–35. https://doi.org/10.1002/jps.24184.

Krebs, M. R.H., E. H.C. Bromley, and A. M. Donald. 2005. “The Binding of Thioflavin-T to Amyloid Fibrils: Localisation and Implications.” *Journal of Structural Biology* 149 (1): 30–37. https://doi.org/10.1016/j.jsb.2004.08.002.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton. 2012. “ImageNet Classification with Deep Convolutional Neural Networks.” *Advances In Neural Information Processing Systems*, 1–9. https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007.

Kueltzo, Lisa A., Wei Wang, Theodore W. Randolph, and John F. Carpenter. 2008. “Effects of Solution Conditions, Processing Parameters,and Container Materials on Aggregation of AMonoclonal Antibody during Freeze–Thawing.” *Journal of Pharmaceutical Sciences* 9 (5): 1801–12. https://doi.org/10.1002/jps.

Kumru, Ozan S., Jun Liu, Junyan A. Ji, Wilson Cheng, Y. John Wang, Tingting Wang, Sangeeta B. Joshi, C. Russel Middaugh, and David B. Volkin. 2012. “Compatibility, Physical Stability, and Characterization of an IgG4 Monoclonal Antibody After Dilution into Different Intravenous Administration Bags.” *Journal of Pharmaceutical Sciences* 101 (10): 3636–50. https://doi.org/10.1002/jps.

Kusumoto, Yoko, Aleksey Lomakin, David B. Teplow, and George B. Benedek. 1998. “Temperature Dependence of Amyloid β-Protein Fibrillization.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (21): 12277–82. https://doi.org/10.1073/pnas.95.21.12277.

Lamoth, F., K. Jaton, G. Prod’hom, L. Senn, J. Bille, T. Calandra, and O. Marchetti. 2010. “Multiplex Blood PCR in Combination with Blood Cultures for Improvement of Microbiological Documentation of Infection in Febrile Neutropenia.” *Journal of Clinical Microbiology* 48 (10): 3510–16. https://doi.org/10.1128/JCM.00147-10.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep Learning.” *Nature* 521: 436–44. https://doi.org/10.1038/nature14539.

Lehman, Everett J., Misty J. Hein, Sherry L. Baron, and Christine M. Gersic. 2012. “Neurodegenerative Causes of Death among Retired National Football League Players.” *Neurology* 79 (6): 1970–74. https://doi.org/10.1212/01.wnl.0000428873.10254.b7.

Liu, Lu, Wei Qi, Daniel K. Schwartz, Theodore W. Randolph, and John F. Carpenter. 2013. “The Effects of Excipients on Protein Aggregation during Agitation: An Interfacial Shear Rheology Study.” *Journal of Pharmaceutical Sciences* 102 (8): 2460–70. https://doi.org/10.1002/jps.23622.

Liu, Lu, Theodore W. Randolph, and John F. Carpenter. 2012. “Particles Shed from Syringe Filters and Their Effects on Agitation-Induced Protein Aggregation.” *Journal of Pharmaceutical Sciences 2012* 101: 2952–59.

Ludwig, D. Brett, John F. Carpenter, Jean-Bernard Hamel, and Theodore W. Randolph. 2010. “Protein Adsorption and Excipient Effects on Kinetic Stability of Silicone Oil Emulsions.”

Journal of Pharmaceutical Sciences 99: 1721–33.

Maddux, Nathaniel R., Austin L. Daniels, and Theodore W. Randolph. 2017. “Microflow Imaging Analyses Reflect Mechanisms of Aggregate Formation: Comparing Protein Particle Data Sets Using the Kullback-Leibler Divergence.” *Journal of Pharmaceutical Sciences* 106 (5): 1239–48. https://doi.org/10.1016/j.xphs.2017.01.030.

Malucchi, S., A. Sala, F. Gilli, R. Bottero, A. Di Sapio, M. Capobianco, and Antonio Bertolotto. 2004. “Neutralizing Antibodies Reduce the Efficacy of BIFN during Treatment of Multiple Sclerosis.” *Neurology* 62 (11): 2031–37. https://doi.org/10.1212/01.WNL.0000129265.73259.9E.

McCradden, Melissa D., and Michael D. Cusimano. 2018. “Concussions in Sledding Sports and the Unrecognized ‘Sled Head’: A Systematic Review.” *Frontiers in Neurology* 9 (SEP): 1–8. https://doi.org/10.3389/fneur.2018.00772.

McKee, Ann C., Thor D. Stein, Christopher J. Nowinski, Robert A. Stern, Daniel H. Daneshvar, Victor E. Alvarez, Hyo Soon Lee, et al. 2013. “The Spectrum of Disease in Chronic Traumatic Encephalopathy.” *Brain* 136 (1): 43–64. https://doi.org/10.1093/brain/aws307.

Moore, W. V., and P. Leppert. 1980. “Role of Aggregated Human Growth-Hormone (Hgh) in Development of Antibodies to Hgh.” *J Clin Endocrinol Metab* 51 (4): 691–97.

Movafaghi, Sanli, Hao Wu, Irene M. Francino Urdániz, David S. Bull, Mary D. Kelly, Theodore W. Randolph, and Andrew P. Goodwin. 2020. “The Effect of Container Surface Passivation on Aggregation of Intravenous Immunoglobulin Induced by Mechanical Shock.” *Biotechnology Journal* 15 (9): 1–9. https://doi.org/10.1002/biot.202000096.

Nagase, Terumasa, Keiichi Iwaya, Yoshiki Iwaki, Fumio Kotake, Ryuji Uchida, Tsunao Oh-I, Hidenori Sekine, et al. 2014. “Insulin-Derived Amyloidosis and Poor Glycemic Control: A Case Series.” *American Journal of Medicine* 127 (5): 450–54. https://doi.org/10.1016/j.amjmed.2013.10.029.

Narhi, Linda O., Vincent Corvari, Dean C. Ripple, Nataliya Afonina, Irene Cecchini, Michael R. Defelippis, Patrick Garidel, et al. 2015. “Subvisible (2-100 Mm) Particle Analysis during Biotherapeutic Drug Product Development: Part 1, Considerations and Strategy.” *Journal of Pharmaceutical Sciences* 104 (6): 1899–1908. https://doi.org/10.1002/jps.24437.

Nayak, Arpan, James Colandene, Victor Bradford, and Melissa Perkins. 2011. “Characterization of Subvisible Particle Formation during the Filling Pump Operation of a Monoclonal Antibody Solution.” *Journal of Pharmaceutical Sciences* 100: 4198–4204.

Ohlsson, Pelle, Klara Petersson, Per Augustsson, and Thomas Laurell. 2018. “Acoustic Impedance Matched Buffers Enable Separation of Bacteria from Blood Cells at High Cell Concentrations.” *Scientific Reports* 8 (1): 1–11. https://doi.org/10.1038/s41598-018-25551-0.

Opota, O., K. Jaton, and G. Greub. 2015. “Microbial Diagnosis of Bloodstream Infection: Towards Molecular Diagnosis Directly from Blood.” *Clinical Microbiology and Infection* 21 (4): 323–31. https://doi.org/10.1016/j.cmi.2015.02.005.

Pardeshi, Neha N., Chen Zhou, Theodore W. Randolph, and John F. Carpenter. 2018. “Protein Nanoparticles Promote Microparticle Formation in Intravenous Immunoglobulin Solutions During Freeze-Thawing and Agitation Stresses.” *Journal of Pharmaceutical Sciences* 107 (7): 1852–57. https://doi.org/10.1016/j.xphs.2018.03.016.

Pitt, William, Mahsa Alizadeh, Ghaleb A. Husseini, Daniel S. McClellan, Clara M. Buchanan, Colin G. Bledsoe, Richard A. Robison, et al. 2016. “Rapid Separation of Bacteria from Blood—Review and Outlook” 32 (4): 823–39. https://doi.org/10.5588/ijtld.16.0716.Isoniazid.

Probst, Christine. 2019. “Characterization of Protein Aggregates, Silicone Oil Droplets, and Protein-Silicone Interactions Using Imaging Flow Cytometry.” *Journal of Pharmaceutical Sciences*, 1–11. https://doi.org/10.1016/j.xphs.2019.05.018.

Prummer, O. 1998. “Endogenous Antibodies Directed against Interferon-Alpha: Paradigm of the Immunogenicity of Cytokines and Its Clinical Impact.” *Med Welt* 49 (6): 267–74.

Randolph, Theodore W., and John F. Carpenter. 2007. “Engineering Challenges of Protein Formulations.” *AIChE Journal* 53 (8): 215–28.

Randolph, Theodore W, Elise Schiltz, Donn Sederstrom, Daniel Steinmann, Olivier Mozziconacci, Erwin Freund, Margaret S Ricci, John F Carpenter, and Corrine S Lengsfeld. 2015. “Do Not Drop : Mechanical Shock in Vials Causes Cavitation , Protein Aggregation , and Particle Formation.” Journal of Pharmaceutical Sciences 104: 602–11. https://doi.org/10.1002/jps.24259.

Ratanji, Kirsty D., Jeremy P. Derrick, Rebecca J. Dearman, and Ian Kimber. 2014. “Immunogenicity of Therapeutic Proteins: Influence of Aggregation.” Journal of Immunotoxicology 11 (2): 99–109. https://doi.org/10.3109/1547691X.2013.821564.

Rhodes, Andrew, Laura E. Evans, Waleed Alhazzani, Mitchell M. Levy, Massimo Antonelli, Ricard Ferrer, Anand Kumar, et al. 2017. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Critical Care Medicine. Vol. 45. https://doi.org/10.1097/CCM.0000000000002255.

Ridker, Paul M, Jean-Claude Tardif, Pierre Amarenco, William Duggan, Robert J. Glynn, J. Wouter Jukema, John J.P. Kastelein, et al. 2017. “Lipid-Reduction Variability and Antidrug-Antibody Formation with Bococizumab.” New England Journal of Medicine 376 (16): 1517–26. https://doi.org/10.1056/nejmoa1614062.

Ring, J., W. Stephan, and W. Brendel. 1979. “Anaphylactoid Reactions to Infusions of Plasma Protein and Human-Serum Albumin - Role of Aggregated Proteins and of Stabilizers Added during Production.” *Clin Allergy* 9 (1): 89–97.

Roberts, Christopher J. 2007. “Non-Native Protein Aggregation Kinetics.” *Biotechnology and Bioengineering* 98 (5): 927–38. https://doi.org/10.1002/bit.21627.

———. 2014a. “Protein Aggregation and Its Impact on Product Quality.” *Current Opinion in Biotechnology* 30: 211–17. https://doi.org/10.1016/j.copbio.2014.08.001.

———. 2014b. “Therapeutic Protein Aggregation: Mechanisms, Design, and Control.” *Trends in Biotechnology* 32 (7): 372–80. https://doi.org/10.1016/j.tibtech.2014.05.005.

Rombach-Riegraf, Verena, Anette C. Karle, Babette Wolf, Laetitia Sordé, Stephan Koepke, Sascha Gottlieb, Jennifer Krieg, et al. 2014. “Aggregation of Human Recombinant Monoclonal Antibodies Influences the Capacity of Dendritic Cells to Stimulate Adaptive T-Cell Responses in Vitro.” *PLoS ONE* 9 (1). https://doi.org/10.1371/journal.pone.0086322.

Rosenberg, Amy S. 2006. “Effects of Protein Aggregates: An Immunologic Perspective.” *The AAPS Journal* 8 (3): E501–7. https://doi.org/10.1208/aapsj080359.

Rosenblatt, Murray. 1952. “Remarks on a Multivariate Transformation.” *The Annals of Mathematical Statistics* 23 (3): 470–72. https://doi.org/10.1214/aoms/1177729394.

Rousseeuw, Peter J. 1987. “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis.” *Journal of Computational and Applied Mathematics* 20 (C): 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

Ryff, Jean Charles, and Huub Schellekens. 2002. “Immunogenicity of RDNA-Derived Pharmaceuticals.” *Trends in Pharmacological Sciences* 23 (6): 254–56. https://doi.org/10.1016/S0165-6147(02)02024-2.

Saggu, Miguel, Ankit R. Patel, and Theodoro Koulis. 2017. “A Random Forest Approach for Counting Silicone Oil Droplets and Protein Particles in Antibody Formulations Using Flow Microscopy.” *Pharmaceutical Research* 34 (2): 479–91. https://doi.org/10.1007/s11095-016-2079-x.

Saller, Verena, Constanze Hediger, Julia Matilainen, Ulla Grauschopf, Karoline Bechtold-Peters, Hanns Christian Mahler, and Wolfgang Friess. 2016. “Influence of Particle Shedding from Silicone Tubing on Antibody Stability.” *Journal of Pharmacy and Pharmacology* 70 (Iii): 675–85. https://doi.org/10.1111/jphp.12603.

Sarciaux, Jeanne Marie, Said Mansour, Michael J. Hageman, and Steven L. Nail. 1999. “Effects of Buffer Composition and Processing Conditions on Aggregation of Bovine IgG during Freeze-Drying.” *Journal of Pharmaceutical Sciences* 88 (12): 1354–61. https://doi.org/10.1021/js980383n.

Sauerborn, Melody, Vera Brinks, Wim Jiskoot, and Huub Schellekens. 2010. “Immunological Mechanism Underlying the Immune Response to Recombinant Human Protein Therapeutics.” *Trends in Pharmacological Sciences* 31 (2): 53–59. https://doi.org/10.1016/j.tips.2009.11.001.

Schellekens, Huub. 2002a. “Bioequivalence and the Immunogenicity of Biopharmaceuticals.” *Nature Reviews Drug Discovery* 1 (6): 457–62. https://doi.org/10.1038/nrd818.

———. 2002b. “Immunogenicity of Therapeutic Proteins: Clinical Implications and Future Prospects.” *Clinical Therapeutics* 24 (11): 1720–40. https://doi.org/10.1016/S0149-2918(02)80075-3.

———. 2003. “Immunogenicity of Therapeutic Proteins.” *Nephrology Dialysis Transplantation* 18 (7): 1257–59. https://doi.org/10.1093/ndt/gfg164.

Schroff, Florian, Dmitry Kalenichenko, and James Philbin. 2015. “FaceNet: A Unified Embedding for Face Recognition and Clustering.” *Proceedings of the IEEE Computer*

Society Conference on Computer Vision and Pattern Recognition 07-12-June: 815–23. https://doi.org/10.1109/CVPR.2015.7298682.

Scott, David W. 2015. Multivariate Density Estimation.

Searles, James A., John F. Carpenter, and Theodore W. Randolph. 2001. “The Ice Nucleation Temperature Determines the Primary Drying Rate of Lyophilization for Samples Frozen on a Temperature-Controlled Shelf.” *Journal of Pharmaceutical Sciences* 90 (7): 860–71. https://doi.org/10.1002/jps.1039.

Sediq, A. S., R. Klem, M. R. Nejadnik, P. Meij, and Wim Jiskoot. 2018. “Label-Free, Flow-Imaging Methods for Determination of Cell Concentration and Viability.” *Pharmaceutical Research* 35 (8). https://doi.org/10.1007/s11095-018-2422-5.

Seefeldt, M. B., M. S. Rosendahl, J. L. Cleland, and L. K. Hesterberg. 2009. “Application of High Hydrostatic Pressure to Dissociate Aggregates and Refold Proteins.” *Current Pharmaceutical Biotechnology* 10 (4): 447–55.

Sethuraman, A, T Morcone, and G Belfort. 2004. “Protein Aggregation at Interfaces.” *Protein Science* 13: 223.

Seymour, Christopher W., Foster Gesten, Hallie C. Prescott, Marcus E. Friedrich, Theodore J. Iwashyna, Gary S. Phillips, Stanley Lemeshow, Tiffany Osborn, Kathleen M. Terry, and Mitchell M. Levy. 2017. “Time to Treatment and Mortality during Mandated Emergency Care for Sepsis.” *New England Journal of Medicine* 376 (23): 2235–44. https://doi.org/10.1056/nejmoa1703058.

Sharma, Ashutosh, Sean Anderson, and Anurag S Rathore. 2008. “Filter Clogging Issues in Sterile Filtration.” *Biopharm International 2008* 21: 53–57.

Sharma, Deepak K., Dave King, Peter Oma, and Clark Merchant. 2010. “Micro-Flow Imaging: Flow Microscopy Applied to Sub-Visible Particulate Analysis in Protein Formulations.” *The AAPS Journal* 12 (3): 455–64. https://doi.org/10.1208/s12248-010-9205-1.

Sharma, Deepak K., Peter Oma, Mark J. Pollo, and Muppalla Sukumar. 2010. “Quantification and Characterization of Subvisible Proteinaceous Particles in Opalescent MAb Formulations Using Micro-Flow Imaging.” *Journal of Pharmaceutical Sciences* 99: 2628–42.

Shomali, Maliheh, Angelika Freitag, Julia Engert, Michael Siedler, Zehra Kaymakcalan, Gerhard Winter, John F. Carpenter, and Theodore W. Randolph. 2014. “Antibody Responses in Mice to Particles Formed from Adsorption of a Murine Monoclonal Antibody onto Glass Microparticles.” *Journal of Pharmaceutical Sciences* 103 (1): 78–89. https://doi.org/10.1002/jps.23772.

Shomali, Maliheh, Sultan Tanriverdi, Angelika J. Freitag, Julia Engert, Gerhard Winter, Michael Siedler, Zehra Kaymakcalan, John F. Carpenter, and Theodore W. Randolph. 2015. “Dose Levels in Particulate-Containing Formulations Impact Anti-Drug Antibody Responses to Murine Monoclonal Antibody in Mice.” *Journal of Pharmaceutical Sciences*. https://doi.org/10.1002/jps.24413.

Sieracki, Christian K., Michael E. Sieracki, and Charles S. Yentsch. 1998. “An Imaging-in-Flow System for Automated Analysis of Marine Microplankton.” *Marine Ecology Progress Series* 168: 285–96. https://doi.org/10.3354/meps168285.

Simonyan, Karen, and Andrew Zisserman. 2015. “Very Deep Convolutional Networks for Large-Scale Image Recognition.” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.

Singh, Satish K. 2011. “Impact of Product-Related Factors on Immunogenicity of Biotherapeutics.” *Journal of Pharmaceutical Sciences* 100 (2): 354–87.

Singh, Satish K., Nataliya Afonina, Michel Awwad, Karoline Bechtold-Peters, Jeffery T. Blue, Danny Chou, Mary Cromwell, et al. 2010. “An Industry Perspective on the Monitoring of Subvisible Particles as a Quality Attribute for Protein Therapeutics.” *Journal of Pharmaceutical Sciences* 99 (8): 3302–21.

Sinha, Mridu, Julietta Jupe, Hannah Mack, Todd P Coleman, Shelley M Lawrence, and I Fraley. 2018. “Emerging Technologies for Molecular Diagnosis of Sepsis.” *Clinical Microbiology Reviews* 31 (2): 1–26.

Siska, Christine, Paul Harber, and Bruce A. Kerwin. 2020. “Shocking Data on Parcel Shipments of Protein Solutions.” *Journal of Pharmaceutical Sciences* 109: 690–95.

Sluzky, Victoria, Alexander M. Klibanov, and Robert Langer. 1992. “Mechanism of Insulin Aggregation and Stabilization in Agitated Aqueous-Solutions.” *Biotechnology and Bioengineering 1992* 40: 895–903.

Snell, Jared R., Connor R. Monticello, Cheng Her, Emma L. Ross, Ashley A. Frazer-Abel, John F. Carpenter, and Theodore W. Randolph. 2020. “DEHP Nanodroplets Leached From Polyvinyl Chloride IV Bags Promote Aggregation of IVIG and Activate Complement in Human Serum.” *Journal of Pharmaceutical Sciences* 109 (1): 429–42. https://doi.org/10.1016/j.xphs.2019.06.015.

Strambini, Giovanni B., and Edi Gabellieri. 1996. “Proteins in Frozen Solutions: Evidence of Ice-Induced Partial Unfolding.” *Biophysical Journal* 70 (2 I): 971–76. https://doi.org/10.1016/S0006-3495(96)79640-6.

Strambini, Giovanni B., and Margherita Gonnelli. 2007. “Protein Stability in Ice.” *Biophysical Journal* 92 (6): 2131–38. https://doi.org/10.1529/biophysj.106.099531.

Strehl, Rene, Verena Rombach-Riegraf, Manuel Diez, Kamal Egodage, Markus Bluemel, Margit Jeschke, and Atanas V. Koulov. 2012. “Discrimination between Silicone Oil Droplets and Protein Aggregates in Biopharmaceuticals: A Novel Multiparametric Image Filter for Sub-Visible Particles in Microflow Imaging Analysis.” *Pharmaceutical Research* 29 (2): 594–602. https://doi.org/10.1007/s11095-011-0590-7.

Sun, Yi, Yuheng Chen, Xiaogang Wang, and Xiaoou Tang. 2014. “Deep Learning Face Representation by Joint Identification-Verification.” *Advances in Neural Information Processing Systems* 3 (January): 1988–96.

Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov,

Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. “Going Deeper with Convolutions.” *IEEE Conference in Computer Vision and Pattern Recognition*. https://doi.org/10.1109/CVPR.2015.7298594.

Taigman, Yaniv, Ming Yang, Marc’Aurelio Ranzato, and Lior Wolf. 2014. “DeepFace: Closing the Gap to Human-Level Performance in Face Verification.” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, January. https://doi.org/10.1109/CVPR.2014.220.

Tamilvanan, Shunmugaperumal, Natarajan Livingston Raja, Biswanath Sa, and Sanat Kumar Basu. 2010. “Clinical Concerns of Immunogenicity Produced at Cellular Levels by Biopharmaceuticals Following Their Parenteral Administration into Human Body.” *Journal of Drug Targeting* 18 (7): 489–98. https://doi.org/10.3109/10611861003649746.

Teska, Brandon M., Jeffrey M. Brake, Gregory S. Tronto, and John F. Carpenter. 2016. “Aggregation and Particle Formation of Therapeutic Proteins in Contact With a Novel Fluoropolymer Surface Versus Siliconized Surfaces: Effects of Agitation in Vials and in Prefilled Syringes.” *Journal of Pharmaceutical Sciences* 105 (7): 2053–65. https://doi.org/10.1016/j.xphs.2016.04.015.

Thirumangalathu, Renuka, Sampathkumar Krishnan, David N. Brems, Theodore W. Randolph, and John F. Carpenter. 2006. “Effects of PH, Temperature, and Sucrose on Benzyl Alcohol-Induced Aggregation of Recombinant Human Granulocyte Colony Stimulating Factor.” *Journal of Pharmaceutical Sciences* 95 (7): 1480–97.

Thorpe, T. C., M. L. Wilson, J. E. Turner, J. L. DiGuiseppi, M. Willert, S. Mirrett, and L. B. Reller. 1990. “BacT/Alert: An Automated Colorimetric Microbial Detection System.” *Journal of Clinical Microbiology* 28 (7): 1608–12.

Torisu, Tetsuo, Takahiro Maruno, Yoshinori Hamaji, Tadayasu Ohkubo, and Susumu Uchiyama. 2017. “Synergistic Effect of Cavitation and Agitation on Protein Aggregation.” *Journal of Pharmaceutical Sciences* 106 (2): 521–29. https://doi.org/10.1016/j.xphs.2016.10.015.

Torosantucci, R., C. Schoneich, and Wim Jiskoot. 2014. “Oxidation of Therapeutic Proteins and Peprides: Structural and Biological Consequences.” *Pharmaceutical Research* 11 (2): 99–109.

Totty, H., M. Ullery, J. Spontak, J. Viray, M. Adamik, B. Katzin, W. M. Dunne, and P. Deol. 2017. “A Controlled Comparison of the BacT/ALERT® 3D and VIRTUO™ Microbial Detection Systems.” *European Journal of Clinical Microbiology and Infectious Diseases* 36 (10): 1795–1800. https://doi.org/10.1007/s10096-017-2994-8.

Twomey, Alan, Rebekah Less, Kosaku Kurata, Hiroshi Takamatsu, and Alptekin Aksan. 2013. “In Situ Spectroscopic Quantification of Protein-Ice Interactions.” *Journal of Physical Chemistry B* 117 (26): 7889–97. https://doi.org/10.1021/jp403267x.

Tyagi, Anil K., Theodore W. Randolph, Aichun Dong, Kevin M. Maloney, Carl Jr. Hiitscherich, and John F. Carpenter. 2009. “IgG Particle Formation during Filling Pump Omperation: A Case Study of Heterogeneous Nucleation on Stainless Steel Nanoparticles.” *Journal of Pharmaceutical Sciences* 98: 94–104. https://doi.org/10.1002/jps.

Tzannis, Stelios T., William J. M. Hrushesky, Patricia A. Wood, and Todd M. Przybycien. 1996.

"Irreversible Inactivation of Interleukin 2 in a Pump-Based Delivery Environment." *Proceedings of the National Academy of Sciences* 93: 5460–65.

"USP<788>, United States Pharmacopeia, USP35-NF30, 20102. Particulate Matter in Injections. United States Pharmacopeial Convention." n.d. In .

"USP<789>, United States Pharmacopeia, USP29-NF24, 2012. Particulate Matter in Opthalmic Solutions. United States Pharmacopeial Convention." n.d.

Vandesteeg, Nathan, and Cathy Kilbert. 2013. "Differentiation of Subvisible Silicone Oil Droplets from Irregular Standard Dust Particles." *Journal of Pharmaceutical Sciences* 10: 1696–1700. https://doi.org/10.1002/jps.

Vincent, Jean Louis, John C. Marshall, S. A. Ñamendys-Silva, Bruno François, Ignacio Martin-Loeches, Jeffrey Lipman, Konrad Reinhart, et al. 2014. "Assessment of the Worldwide Burden of Critical Illness: The Intensive Care Over Nations (ICON) Audit." *The Lancet Respiratory Medicine* 2 (5): 380–86. https://doi.org/10.1016/S2213-2600(14)70061-X.

Vlieland, N. D., M. R. Nejadnik, H. Gardarsdottir, S. Romeijn, A. S. Sediq, M. L. Bouvy, A. C.G. Egberts, B. J.F. van den Bemt, and W. Jiskoot. 2018. "The Impact of Inadequate Temperature Storage Conditions on Aggregate and Particle Formation in Drugs Containing Tumor Necrosis Factor-Alpha Inhibitors." *Pharmaceutical Research* 35 (2): 1–11. https://doi.org/10.1007/s11095-017-2341-x.

Vörös, Janos. 2004. "The Density and Refractive Index of Adsorbing Protein Layers." *Biophysical Journal* 87 (1): 553–61. https://doi.org/10.1529/biophysj.103.030072.

Wain, John, To Song Diep, Vo Anh Ho, Amanda M Walsh, Nguyen Thi, Tuyet Hoa, Christopher M Parry, J Nicholas, T O Song Diep, and V O a N H Ho. 1998. "Quantitation of Bacteria in Blood of Typhoid Fever Patients and Relationship between Counts and Clinical Features , Transmissibility , and Antibiotic Resistance Quantitation of Bacteria in Blood of Typhoid Fever Patients and Relationship between Counts An" 36 (6): 1683–87.

Walker, Kendall R., and Giuseppina Tesco. 2013. "Molecular Mechanisms of Cognitive Dysfunction Following Traumatic Brain Injury." *Frontiers in Aging Neuroscience* 5 (JUL): 1–25. https://doi.org/10.3389/fnagi.2013.00029.

Walsh, Gary. 2018. "Biopharmaceutical Benchmarks 2018." *Nature Biotechnology* 36 (12): 1136–45. https://doi.org/10.1038/nbt.4305.

Wang, Chen, Xiao Zhong, David B. Ruffner, Alexandra Stutt, Laura A. Philips, Michael D. Ward, and David G. Grier. 2016. "Holographic Characterization of Protein Aggregates." *Journal of Pharmaceutical Sciences* 105 (3): 1074–85. https://doi.org/10.1016/j.xphs.2015.12.018.

Wang, Haohan, Xindi Wu, Zeyi Huang, and Eric P. Xing. 2019. "High Frequency Component Helps Explain the Generalization of Convolutional Neural Networks," no. Remark 1: 8684–94. http://arxiv.org/abs/1905.13545.

Wang, Liwei, Yin Li, Jing Huang, and Svetlana Lazebnik. 2018. "Learning Two-Branch Neural Networks for Image-Text Matching Tasks." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–14. https://doi.org/10.1109/TPAMI.2018.2797921.

Wang, Wei. 1999. Instability, Stabilization, and Formulation of Liquid Protein Pharmaceuticals. International Journal of Pharmaceutics. Vol. 185.

Wang, Wei, and Christopher J. Roberts. 2018. “Protein Aggregation – Mechanisms, Detection, and Control.” *International Journal of Pharmaceutics* 550 (1–2): 251–68. https://doi.org/10.1016/j.ijpharm.2018.08.043.

Wang, Wei, Satish K. Singh, Ning Li, Maria R. Toler, Kevin R. King, and Sandeep Nema. 2012. “Immunogenicity of Protein Aggregates - Concerns and Realities.” *International Journal of Pharmaceutics* 431 (1–2): 1–11. https://doi.org/10.1016/j.ijpharm.2012.04.040.

Wasylaschuk, Walter R., Paul A. Harmon, Gabriella Wagner, Amy B. Harman, Allen C. Templeton, Hui Xu, and Robert A. Reed. 2007. “Evaluation of Hydroperoxides in Common Pharmaceutical Excipients.” *Journal Of Pharmaceutical Sciences 2007* 96: 106–16.

Webb, S., J. L. Cleland, John F. Carpenter, and Theodore W. Randolph. 2002. “Protein Aggregation at Interfaces Formed during Lyophilization.” *Abstracts of Papers of the American Chemical Society* 224: 163-BIOT.

Westh, Henrik, G. Lisby, F. Breysse, B. Böddinghaus, M. Chomarat, V. Gant, A. Goglio, et al. 2009. “Multiplex Real-Time PCR and Blood Culture for Identification of Bloodstream Pathogens in Patients with Suspected Sepsis.” *Clinical Microbiology and Infection* 15 (6): 544–51. https://doi.org/10.1111/j.1469-0691.2009.02736.x.

Wu, Hao, Carly F. Chisholm, Meagen Puryear, Sanli Movafaghi, Samuel D. Smith, Yekaterina Pokhilchuk, Corinne S. Lengsfeld, and Theodore W. Randolph. 2020. “Container Surfaces Control Initiation of Cavitation and Resulting Particle Formation in Protein Formulations After Application of Mechanical Shock.” *Journal of Pharmaceutical Sciences* 109 (3): 1270–80. https://doi.org/10.1016/j.xphs.2019.11.015.

Wu, Hao, and Theodore W. Randolph. 2019. “Rapid Quantification of Protein Particles in High-Concentration Antibody Formulations.” *Journal of Pharmaceutical Sciences* 108 (3): 1110–16. https://doi.org/10.1016/j.xphs.2018.10.021.

———. 2020. “Aggregation and Particle Formation During Pumping of an Antibody Formulation Are Controlled by Electrostatic Interactions Between Pump Surfaces and Protein Molecules.” *Journal of Pharmaceutical Sciences* 109 (4): 1473–82. https://doi.org/10.1016/j.xphs.2020.01.023.

Wynn, James L. 2016. “Defining Neonatal Sepsis.” *Curr Opin Paediatr* 28 (2): 135–40. https://doi.org/10.1097/MOP.0000000000000315.Defining.

Zölls, Sarah, Manuel Gregoritza, Ruedeeporn Tantipolphan, Michael Wiggenhorn, Gerhard Winter, Wolfgang Friess, and Andrea Hawe. 2013. “How Subvisible Particles Become Invisible — Relevance of the Refractive Index for Protein Particle Analysis.” *Pharmaceutical Biotechnology* 102 (5): 1434–46. https://doi.org/10.1002/jps.

Zölls, Sarah, Daniel Weinbuch, Michael Wiggenhorn, Gerhard Winter, Wolfgang Friess, Wim Jiskoot, and Andrea Hawe. 2013. “Flow Imaging Microscopy for Protein Particle Analysis-A Comparative Evaluation of Four Different Analytical Instruments.” *The AAPS Journal* 15 (4): 1200–1211. https://doi.org/10.1208/s12248-013-9522-2.

Appendix A: Algorithm Implementation Details

Many of the ConvNets trained throughout this thesis were trained using a triplet loss approach borrowed from facial recognition[1]. These algorithms are used to train ConvNets that accept images and return low-dimensional representations of these images. These networks were trained to learn and return very low-dimensional (i.e. 2D) representations of FIM images for easier analysis. This appendix describes how this approach was implemented for most of the analyses described in this thesis in more detail. Additionally, this section also describes some optimization that was used when designing the alternative method for performing the container study described in the fibrillation analysis described in Chapter 6.

A.1 Triplet Loss and Mining

In a triplet loss training scheme, the low-dimensional representations or embeddings returned by the ConvNet are optimized by adjusting network parameters so that these representations minimize a triplet loss function. The loss function is evaluated on triplets or sets of three images consisting of one image of a specific particle type (the anchor image), a second image of the same particle type (the positive image), and an image of a different particle type (the negative image). The triplet loss L is typically computed using:

$$L = \max\left(d_{an} - d_{ap} - \alpha, 0\right) \tag{1}$$

where d_{ap} is the Euclidean distance between the anchor and positive image embeddings, d_{an} is the Euclidean distance between the anchor and negative image embeddings, and α is the margin, a small constant that sets how much further we wanted d_{an}relative to d_{ap}. This loss is zero if the distance between the anchor and negative image embeddings is at least a distance α greater than that between the anchor and positive image embeddings. Note that other versions of this triplet loss such as the modified triplet loss used in Chapter 4 have also been proposed and used.

Triplets can be divided into three general types depending on their loss function value and the relative values of d_{ap} and d_{an}[1]. An “easy” triplet is one for which the anchor-negative distance is at least a value α greater than the anchor-positive distance, resulting in zero loss. A “hard” triplet is one with embeddings so that the anchor image is closer to the negative image than the positive image in the embedding space ($d_{ap} > d_{an}$). A “semihard” triplet is a triplet between these extremes: the negative image is farther away from the anchor than the positive image in the embedding space but by less than a distance α and thus has non-zero loss. The “difficulty” and thus the loss of a triplet can be thought of as how easy it is to tell that the anchor image contains the same type of object in the positive image and a different type than that in the negative image. As the loss of a triplet increases, the anchor image increasingly resembles the negative image more than it does the positive image.

During training, triplets are often mined to remove easy and/or hard triplets[1]. Easy triplets that can already be separated correctly by the ConvNet do not contribute to the loss function and, as

they make up an increasingly large fraction of the triplets that can be constructed from the training set as the network trains, can “drown out” triplets that still have a positive loss. Hard triplets meanwhile dominate the loss function and, as it can be difficult to correctly sort these images in the embedding space especially during the early part of training, can cause the network to collapse on trivial solutions (e.g. ConvNet parameters that map all images to a single point in the embedding space). Filtering out these easy and hard triplets can improve the convergence of the model during training. Additionally, as only a subset of all available triplets are used to train the model, this mining introduces randomness during training that can help prevent overfitting. However, as the exact strategy used for mining triplets and selection can also impact the performance of the final trained model[2] it is important to optimize the mining strategy used for a given application of the triplet loss algorithm.

A.2 Triplet Loss Implementations

Different implementations of the triplet loss were used for the various projects in this thesis. These implementations varied in triplet mining strategy and the implementation of the triplet loss section. An initial implementation was used to compute the results in Chapters 4, 5, and 7 and an updated implementation was used in the results in Chapters 6. This section describes these implementations in detail while later sections in this chapter will describe the optimization that went into the updated implementation.

A.2.1 Initial Triplet Loss Implementation

In the initial implementation of the triplet loss, ConvNets were trained using minibatches of semihard triplets. Triplets were assembled by dividing the training images into minibatches of images which were to be used as anchor images for triplets. Each anchor image was only used once per minibatch. To mine triplets for the anchor images, up to 100 candidate positive-negative image pairs were selected from all training images and the anchor-positive and anchor-negative distances were computed for each using the embeddings for each image. These embeddings were typically computed at the start of each epoch and were used for triplet mining through the end of the epoch—an offline triplet mining strategy. A random semihard triplet among these candidates was then selected and used to train the algorithm in that minibatch. If no semihard triplets were mined one of the candidate positive-negative image pairs was randomly selected to construct the triplet.

Once a minibatch of triplets was selected, image augmentation was applied to each image before use in training. This augmentation was designed both to reduce the risk of overfitting and to help mask common FIM image artifacts that may otherwise be used when computing embeddings. This augmentation included global image noise via addition of a random RGB value to all pixels in an image and flipping the image horizontally and vertically. Image noise was added to introduce random changes in image color to help mask changes in background image color that can often vary between FIM measurements. These adjustments were performed by sampling a random RGB value from a normal distribution with zero mean and 0.01 standard deviation and adding that value to each pixel in the image. Flipping was included to help ensure the algorithm was robust against both the orientation of the particle in the image and any imperfections in the particle illumination by the instrument. Images were flipped 50% of

the time in each dimension. It is important to note that this augmentation was performed after triplet mining and thus did not influence if a triplet met the mining criteria.

Minibatches assembled as described above were then used to train the algorithm to minimize the modified triplet loss in Chapter 4 until all minibatches of anchor images were used. At the end of each epoch, the embeddings for the training images were updated, the training images were shuffled into new minibatches and the triplet mining process was repeated. During training, the triplet loss over a small validation set of images was computed at the end of each epoch. This loss was computed using triplets mined in a similar fashion as triplets used in algorithm training but including easy and hard triplets. The ConvNet parameters that resulted in the lowest value of the triplet loss over the validation set was saved and used in subsequent analyses.

While the approach described above was effective at differentiating images from different classes, it differs dramatically from existing implementations of the triplet loss that are better characterized on facial recognition tasks[1,2]. After performing the optimization described later in this section, this initial triplet loss implementation was abandoned in favor of a better-characterized and better-performing implementation.

A.2.2 Batch All Triplet Loss Implementation

The second triplet loss implementation was based heavily on the “Batch All (Nonzero)” strategy described by Hermans et al..[2] In this approach, minibatches of triplets were generated by randomly selecting a small number of images from a few different samples and assembling all valid triplets with a positive triplet loss (i.e. all semihard and hard triplets) that can be constructed from these images. These triplets were then used to train the ConvNet without

applying any image augmentation. At the end of each epoch, the triplet loss on the validation set was computed and the model parameters that minimized this loss was used in later analyses.

A.3 Triplet Loss Optimization

The performance of various triplet loss algorithms on a sample FIM image dataset was compared to develop the updated triplet loss implementation. This FIM image dataset consisted of aggregates of a monoclonal antibody (mAb) formulation donated by AstraZeneca generated by a simulated shipping stress. To generate aggregates, mAb formulation was filled into one of four container types: Glass vials from Schott (Mainz, Germany), high-density polyethylene (HPDE) containers from Meissner (Cryovault; Camarillo, CA), low-density polyethylene (ULDPE) bags from Millipore Sigma (Mobius; Burlington, MA), and Ethylene Vinyl Acetate Monomaterial (EVAM) bags from Sartorius Stedim (CelsiusPak; Göttingen, Germany). Each container was filled to obtain roughly the same air-water interface area-to-volume ratio of fluid in the container. The filled containers were then agitated on a transportation simulator instrument from Lansmont Corporation (Monterey, CA) to simulate 6 hours of air transportation followed by 6 hours of truck transportation. The resulting aggregates were then imaged using a FlowCam® VS instrument (Yokogawa Fluid Imaging Technologies, Inc., Scarborough, ME). 8,000 FIM images per container were designated as training data for the ConvNets with an additional 1,000 FIM images per container used as a validation set. All remaining FIM images (>1,000 per container) were used as a test set for the algorithm.

Table 10: ConvNet structure used in triplet loss optimization

Layer #	Layer Type	# of features	Feature Size	Activation	Input Shape	Output Shape
1	Convolutional	32	3x3	ReLU	24x24x3	22x22x32
2	Convolutional	32	3x3	ReLU	22x22x32	20x20x32
3	Max Pooling (2x2)	-	-	-	20x20x32	10x10x32
4	Convolutional	32	3x3	ReLU	10x10x32	8x8x32
5	Flatten	-	-	-	8x8x32	2,048
6	Dense	64	-	Softplus	2,048	64
7	Dense	2	-	Softmax	64	2

Several ConvNets were trained using a triplet loss approach to differentiate mAb aggregates generated by agitation in the four different containers (Glass, EVAM, HDPE, and ULDPE). The structure of these ConvNets is shown in Table 10. Unless otherwise noted, all ConvNets were trained to minimize the standard triplet loss (eq. 1) with a 0.5 margin, an Adam optimizer at default parameters and no image augmentation. The triplet loss on the validation set was used as early stopping criteria for the training and network training was stopped after 5 epochs without improvement to this triplet loss.

The performance of the different algorithms was compared by visual inspection of the raw embeddings as well as using the goodness-of-fit algorithm. The embeddings yielded by each algorithm were visually inspected to determine if and how the different mining strategies impacted the separation between aggregates from each container. Goodness-of-fit hypothesis testing was used to investigate how these different algorithms influenced the Type I error rate of the hypothesis test on sets of FIM images from the test set for each container. For each test the null PDF was estimated using a kernel density estimate with normal product kernels and bandwidths computed via a normal reference rule. The Type I error rate was computed on 10,000 sets of 50 image embeddings sampled from the test set.

The following sections describe the optimizations that were considered and the optimized settings.

A.3.1 Triplet Mining

We first investigated different strategies for selecting triplets for use in ConvNet training. Each of the investigated strategies was based on either our initial triplet mining implementation or the batch all (nonzero) mining strategy described in the updated implementation section. As these strategies vary both how triplets were mined and the difficulty of the selected triplets, we also investigated several variations on these strategies that are more intermediate between the behavior of the two algorithms. For our “Initial” triplet mining strategy, we investigated the performance of this strategy using both the “Offline” triplet mining described previously as well as an “Online” triplet mining strategy in which the embeddings used in triplet mining are computed after each minibatch rather than after each epoch. The more frequent embedding updates results in triplets that always meet the mining criteria at the current parameters of the ConvNet. For the “Batch All” triplet mining strategy we investigated the performance of the strategy with different strategies for triplet selection. Triplet selection was performed to either select either all non-easy triplets (“Batch All (Nonzero)”) or semihard triplets only as was done in the Initial implementation (“Batch All (Semihard)”). We also investigated the performance of this approach without any triplet selection (“Batch All”) in which all valid triplets were used in training. All triplet selection in the Batch All implementations was done in an online fashion. Networks trained using the Initial implementation were trained on minibatches of 64 semihard triplets. To train the Batch All methods, 16 random images were selected from each of the four containers the algorithm was trained on and all valid triplets that could be constructed from this 64-image set that met the desired triplet selection criteria were used in training.

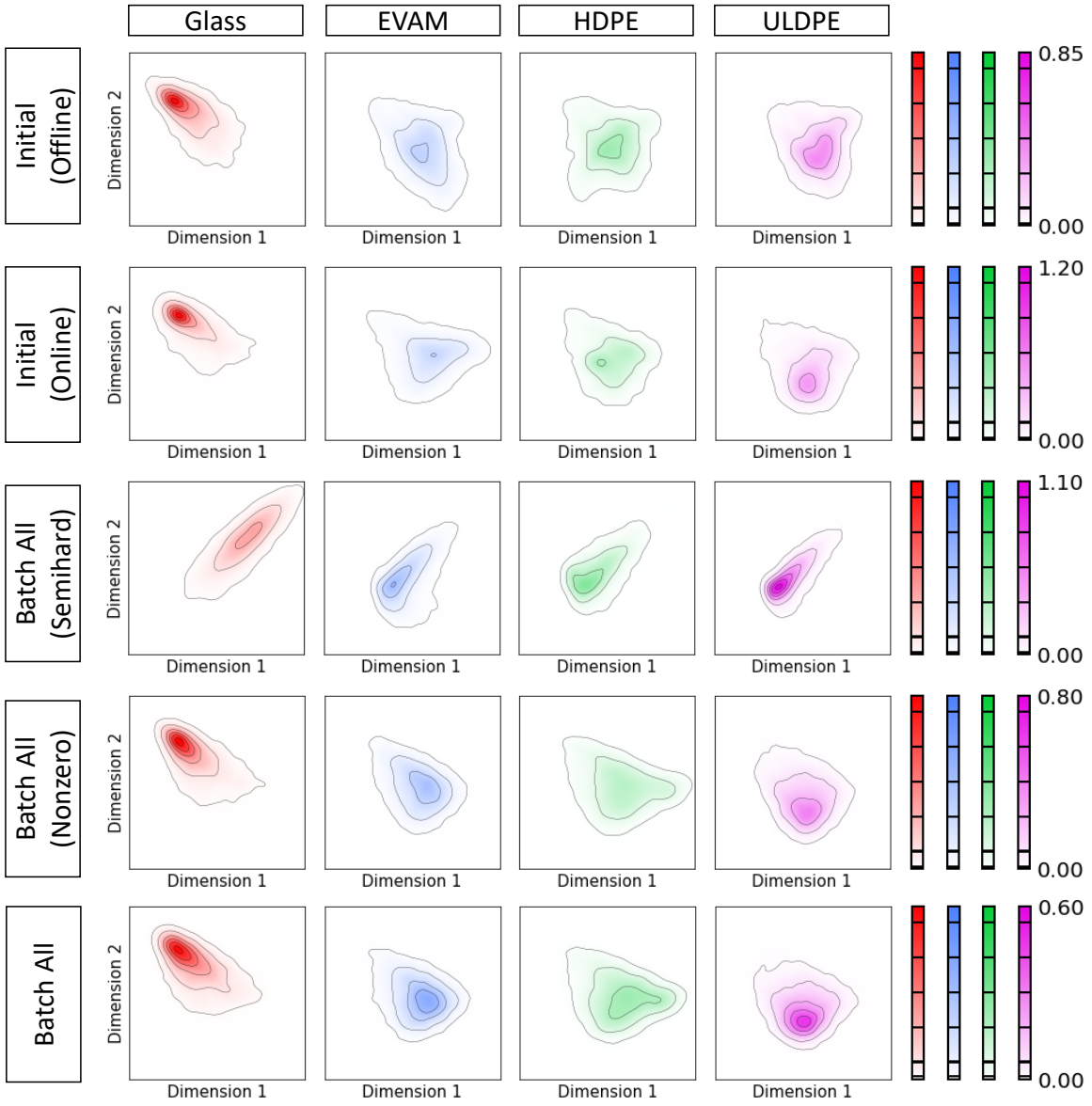

Figure 23: ConvNet Embeddings Using Different Triplet Mining Strategies
Contour plots of FIM image embeddings from ConvNets trained with different triplet mining schemes. Rows correspond to the different triplet mining scheme while columns indicate which container the FIM images used to obtain the embeddings were taken from. The colorbars to the right of each row indicate the relative fraction of embeddings found at each location in the embedding space for that row.

Table 11: Type I Error Rate Using Different Triplet Mining Strategies
Rows indicate ConvNet models trained via different mining strategies, columns indicate particles from the four containers the algorithm was trained to differentiate, values indicate the Type I error rate of the goodness-of-fit hypothesis test on particles from each container by each network.

Mining Strategy	Glass	EVAM	HDPE	ULDPE	Average
Initial (Offline)	9.1%	5.2%	5.3%	5.8%	6.3%
Initial (Online)	8.6%	7.1%	5.2%	8.2%	7.3%
Batch All (Semihard)	4.9%	4.8%	4.9%	5.1%	4.9%
Batch All (Nonzero)	4.7%	5.1%	5.7%	6.5%	5.5%
Batch All	5.4%	5.2%	5.5%	5.3%	5.4%

Figure 23 shows sample contour plots of the resulting embeddings from each container while Table 11 shows the Type I error rates for these networks. As shown in Figure 23, each plot generally yielded similar embedding profiles for the four different containers; Glass generally contained a distinct particle type while EVAM, HDPE, and HDPE yielded similar particle types but in different ratios. While the extent of separability between particles found in the remaining three containers varied, ULDPE was generally enriched in one particle type while EVAM and HDPE were both enriched in a different particle type. The Initial (Online), Batch All, and Batch All (Nonzero) methods offered the best separation between the four different containers. However, the results in Table 11 indicate that the Batch All strategies achieved lower Type I error rates than the Initial method did on this dataset, network architecture, and training setup. Based on both metrics the Batch All and Batch All (Nonzero) methods were the most effective for analyzing these FIM images. The Batch All (Nonzero) method was deemed the superior method due to the slightly improved separation between particles from the four containers and the slightly more stable convergence due to the removal of easy triplets during training. However, the Batch All method is also viable for this analysis due to its similar performance but without using triplet selection.

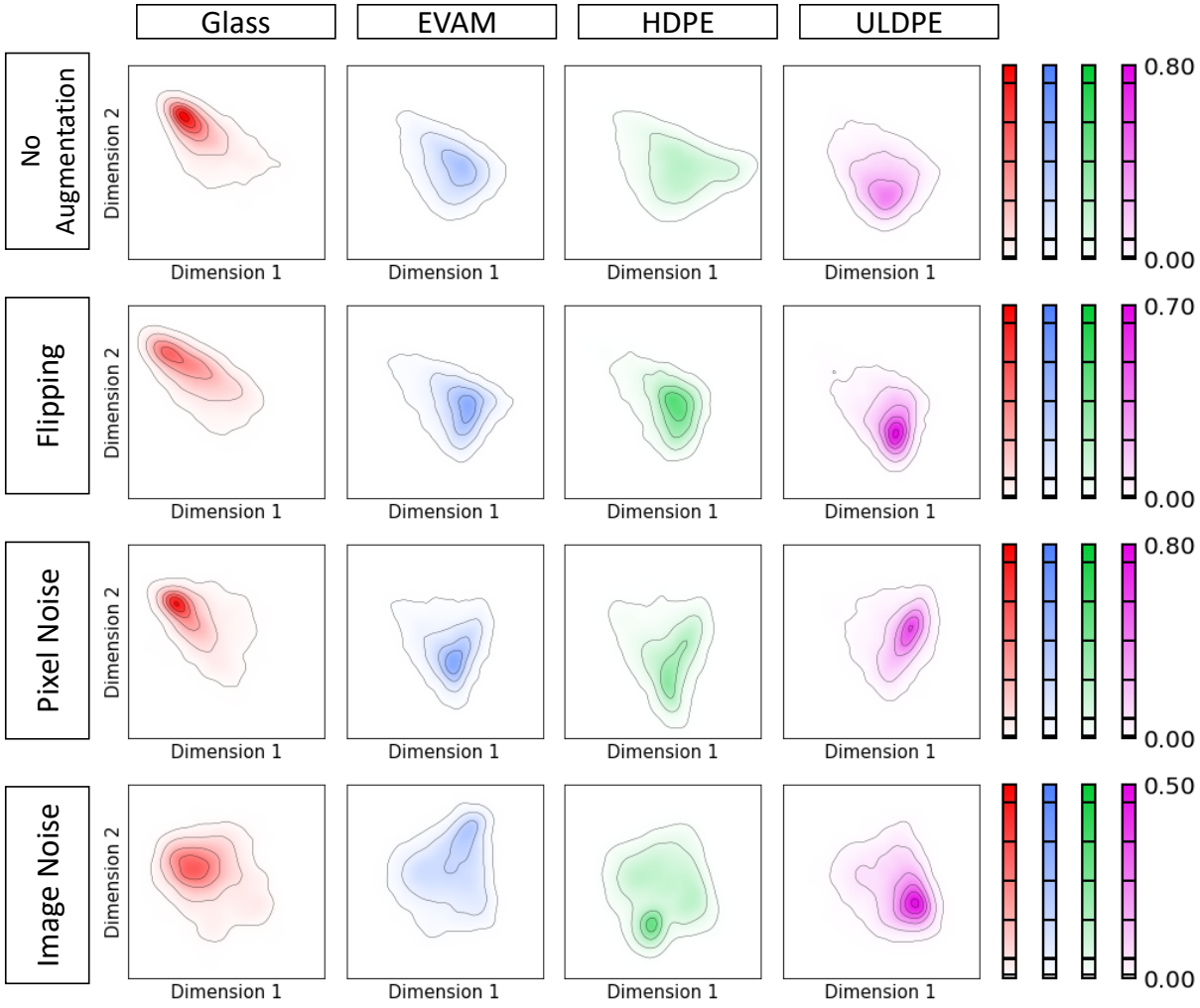

Figure 24: ConvNet Embeddings Using Different Image Augmentation Strategies Contour plots of FIM image embeddings from ConvNets trained with different image augmentations enabled. Rows correspond to the different image augmentations while columns indicate which container (Glass or HDPE) the FIM images used to obtain the embeddings were taken from. The format of the subplots matches that in Figure [1]

A.3.2 Image Augmentation

To see what impact augmentation had on algorithm performance, ConvNets were trained with and without various image augmentation using the Batch All (Nonzero) method and a 0.5 margin. In addition to the image noise and flipping augmentations used in our initial

implementation, we also investigated the performance of an additional pixel noise augmentation which consisted of adding a unique random RGB value sampled from a normal distribution with zero mean and 0.01 standard deviation to each pixel. These augmentations were applied both to the training data during training as well as to the validation and test data when evaluating the performance of the model. The impact of these augmentations was assessed by comparing the shape of the embeddings obtained with and without each augmentation to ensure that the augmentation did not significantly bias the embeddings returned by the model.

As is shown in the contour plots in Figure 24, the flipping and pixel noise augmentations did not appreciably change the shape of the image embeddings from each container. However, the image noise augmentation drastically changes the shape of the embeddings, resolving an additional particle population present in HDPE containers that was less appreciably resolved without this augmentation. The images that are embedded in this new region consist of dark, circular particles that were commonly found in the HDPE sample. The image noise augmentation appeared to bias the embeddings returned by the model and thus should not be used to train these algorithms. For sake of simplicity, augmentation was not used in the models trained later in this section as well as in Chapter 6 of the thesis where this method was used. However, the addition of white noise to the image and flipping can safely be included in the analysis if it is desired to use image augmentation.

Table 12: Total Cluster Variance Using Different Margins

Margin	Glass	EVAM	HDPE	ULDPE	Average
0.10	5.9%	5.9%	5.9%	6.3%	6.0%
0.25	5.4%	5.2%	5.5%	5.3%	5.4%
0.50	4.7%	5.1%	5.7%	6.5%	5.5%
1.00	5.4%	5.2%	5.5%	5.3%	5.4%
1.50	5.4%	5.2%	5.5%	5.3%	5.4%

A.3.3 Margin

The margin used in the triplet loss sets the minimum distance that the negative image in a triplet must be further than the positive image for the triplet to exhibit zero loss. To investigate the impact of this parameter on the performance of the network, ConvNets were trained using five different values of the margin: 0.1, 0.25, 0.5, 1 and 1.5 using the Batch All (Nonzero) approach and no augmentation. 0.5 was the margin value used in the networks trained elsewhere in this thesis. Table 12 shows the Type I error rates for these different models. Similar rejection rates were obtained regardless of the value of the margin. Each model also yielded qualitatively similar embeddings for each sample (data not shown). Any value of the margin can therefore be used without appreciably changing the performance of the algorithm

A.3.4 Conclusions

The studies above were performed to compare and optimize the performance of different triplet loss setups on FIM image embeddings. The Batch All (Nonzero) triplet mining strategy was determined to be the optimal strategy due to its ability to resolve particles from the four different container types and its low Type I error rate relative to the other investigated strategies. Augmentation was also excluded in the optimal algorithm settings as, while two of the three

augmentations investigated here can safely be included in during training, excluding augmentation simplifies algorithm training and the inherent randomness of triplet mining combined with aggressive early stopping criteria likely offers sufficient protection against overfitting. While this model was deemed to be optimal, it is noted that the final embeddings obtained from this model were relatively robust to most changes in the mining strategy, augmentation and the margin. While these settings yielded Type I error rates, many different implementations of the triplet loss algorithm can be used without significantly changing the performance of the algorithm or the conclusions drawn from the analysis.

A.4 References

1. Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2015;07-12-June:815-823. doi:10.1109/CVPR.2015.7298682
2. Hermans A, Beyer L, Leibe B. In Defense of the Triplet Loss for Person Re-Identification. 2017. http://arxiv.org/abs/1703.07737.