frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 January 2022
doi: 10.3389/fmicb.2021.785182

Experimental Designs to Study the Aggregation and Colonization of Biofilms by Video Microscopy With Statistical Confidence

Brian A. Pettygrove[1,2], *Heidi J. Smith*[1,2], *Kyler B. Pallister*[2], *Jovanka M. Voyich*[2], *Philip S. Stewart*[1,3] *and Albert E. Parker*[1,4*]

[1] Center for Biofilm Engineering, Montana State University, Bozeman, MT, United States, [2] Department of Microbiology and Cell Biology, Montana State University, Bozeman, MT, United States, [3] Department of Chemical and Biological Engineering, Montana State University, Bozeman, MT, United States, [4] Department of Mathematical Sciences, Montana State University, Bozeman, MT, United States

OPEN ACCESS

Edited by:
Giovanni Di Bonaventura, University of Studies G. d'Annunzio Chieti and Pescara, Italy

Reviewed by:
Erin Samantha Gloag, The Ohio State University, United States
Derek Fleming, Mayo Clinic, United States
Mathias Müsken, Helmholtz Association of German Research Centers (HZ), Germany

***Correspondence:**
Albert E. Parker
parker@math.montana.edu

Specialty section:
This article was submitted to Infectious Agents and Disease, a section of the journal Frontiers in Microbiology

Received: 28 September 2021
Accepted: 06 December 2021
Published: 13 January 2022

Citation:
Pettygrove BA, Smith HJ, Pallister KB, Voyich JM, Stewart PS and Parker AE (2022) Experimental Designs to Study the Aggregation and Colonization of Biofilms by Video Microscopy With Statistical Confidence. Front. Microbiol. 12:785182. doi: 10.3389/fmicb.2021.785182

The goal of this study was to quantify the variability of confocal laser scanning microscopy (CLSM) time-lapse images of early colonizing biofilms to aid in the design of future imaging experiments. To accomplish this a large imaging dataset consisting of 16 independent CLSM microscopy experiments was leveraged. These experiments were designed to study interactions between human neutrophils and single cells or aggregates of *Staphylococcus aureus* (*S. aureus*) during the initial stages of biofilm formation. Results suggest that in untreated control experiments, variability differed substantially between growth phases (i.e., lag or exponential). When studying the effect of an antimicrobial treatment (in this case, neutrophil challenge), regardless of the inoculation level or of growth phase, variability changed as a frown-shaped function of treatment efficacy (i.e., the reduction in biofilm surface coverage). These findings were used to predict the best experimental designs for future imaging studies of early biofilms by considering differing (i) numbers of independent experiments; (ii) numbers of fields of view (FOV) per experiment; and (iii) frame capture rates per hour. A spreadsheet capable of assessing any user-specified design is included that requires the expected mean log reduction and variance components from user-generated experimental results. The methodology outlined in this study can assist researchers in designing their CLSM studies of antimicrobial treatments with a high level of statistical confidence.

Keywords: microscopy, biofilm, image analysis, statistical confidence, repeatability, experimental design

INTRODUCTION

Biofilms consist of matrix embedded aggregates of microorganisms, are widespread in natural and engineered environments, and are physiologically distinct from free-living communities. Biofilm development begins with the loose association of microbes, most often on a surface, followed by a more robust cellular adhesion and aggregation (Hall-Stoodley et al., 2004). The process of biofilm development is an active area of research that has been advanced by modeling, imaging,

and "omic" based approaches (de Beer et al., 1994; Sauer et al., 2002; Rani et al., 2005; Connolly et al., 2015). One of the most important developments in the study of biofilm formation, structure, and dynamics has been the application of optical imaging, specifically confocal laser scanning microscopy (CLSM). CLSM has the ability to provide 3D information of hydrated intact biofilms non-invasively and in real-time (Stewart et al., 2009; Davison et al., 2010; Franklin et al., 2015). CLSM studies have provided valuable insight into biofilms including localization of gene expression (McLoon et al., 2011; Hung et al., 2013), analysis of extracellular material (Allesen-Holm et al., 2006; Baird et al., 2012; Gloag et al., 2013), community organization (Brileya et al., 2014), flow patterns (Martínez-García et al., 2018), the diffusivity of solutes (Stewart et al., 2009), and the spatio-temporal patterns of biocide action (Bridier et al., 2011), making it a powerful tool for any biofilm researcher. However, when studying developing biofilms, the potential for increased variability within early stage biofilms in combination with phototoxicity increases the need for the development of methodological approaches that carefully consider all steps of an imaging experiment including experimental design, data collection, and the reporting of results (Lee and Kitaoka, 2018).

It is standard practice in biological research to establish the repeatability and reproducibility of a particular technique or method and this trend has extended into the biofilm field. There is increasing recognition of the importance of standardized methods in biofilm research to ensure reproducibility across labs (Malone et al., 2017). Efforts have been made to develop standard methods for several biofilm assays (Gomes et al., 2018) such as the drip-flow reactor (Goeres et al., 2009, 2020), single-tube method (Goeres et al., 2019), the CDC biofilm reactor (Goeres et al., 2005), and the 96 well MBEC assay (Parker et al., 2014). However, studies recommending standard practices for imaging-based biofilm experiments have not been similarly reported. A comprehensive review by Magana et al. (2018) found that biofilm methodology was under-represented as a percentage of all published papers in the field with only 4.9% of papers focusing on assay or method development. Conversely, for the field of genome editing, they found 23.4% of papers to be related to methodology. We have searched the biofilm literature for papers on the topic of biofilm imaging methods and observed that the majority of biofilm imaging methods papers are dedicated novel image analysis techniques and packages that assist researchers in quantifying biofilm characteristics such as volume, structure, and surface coverage (Heydorn et al., 2000; Yang et al., 2000; Daims et al., 2006; Mueller et al., 2006; Yerly et al., 2007; Milferstedt et al., 2009; Renslow et al., 2011; Almstrand et al., 2013; Sommerfeld Ross et al., 2014; Tolker-Nielsen and Sternberg, 2014; Larimer et al., 2016; Vyas et al., 2016; Baudin et al., 2017; Luo et al., 2018, 2019; Parker et al., 2018b, 2020). While these studies are essential to further the field of biofilm research by improving the quality and accuracy of data acquired through imaging, we are aware of few papers that include any recommendations for the design of biofilm imaging experiments to assure that data are representative, repeatable across experiments, and reproducible across labs (Korber et al., 1993; Lawrence and Neu, 1999; Heydorn et al., 2000; Daims and Wagner, 2011; Menzel et al., 2016). To answer questions of interest from imaging data with a high level of statistical confidence, a quantitative assessment of variability in biofilm imaging data that specifies the requisite number of experiments and fields of view (FOV) is necessary but currently is lacking. To achieve statistical confidence, there is often an inclination to collect as many FOVs as possible. This approach has the potential to negatively impact data quality in several ways. Excessive imaging of a sample can result in photosensitivity of fluorophores and decrease data quality, although next generation imaging technologies such as light-sheet microscopy are beginning to mitigate these concerns somewhat (Power and Huisken, 2017; Qin et al., 2020). Increasing the number of FOVs being imaged can decrease the temporal resolution for time lapse microscopy, which can be a large detriment when studying rapid processes. Furthermore, advancements in imaging technology which have allowed researchers to generate massive amounts of data by imaging many FOVs over high temporal resolutions require a significant amount of time and computational resources to process and analyze collected images. Thus, identifying the correct balance between sample number and statistical confidence can make imaging-based experiments more efficient while maintaining high quality data and providing confidence in the downstream interpretation.

This paper provides guidance on designing experiments and interpreting resulting data from time-lapse visualization studies with an emphasis on early colonization biofilms. Recommendations are given for the number of experiments, FOVs and frames per hour that should be collected for imaging studies. Data were mined from 16 independent CLSM microscopy experiments that were designed to study interactions between human neutrophils (polymorphonuclear leukocytes or PMNs) and single cells or aggregates of *S. aureus* during the initial stages of biofilm formation. These data, including biological replicates (i.e., independent experiments) and varying inoculum concentrations, were then analyzed to identify the variability associated with the different levels of replication. We use the results to suggest future microscopy experimental designs. Based on the variability identified with each level of replication, results presented herein allowed us to recommend how often to take CLSM images (i.e., the temporal resolution of the sequence of images), how many FOVs to include, and how many replicate experiments to perform when studying initial attachment of biofilms using surface coverage as a quantitative response. While the presented analysis and results are specific to the data set described above, any pilot data can readily be analyzed using the approach that we present to inform experimental designs. We include our experimental design assessment tool as an Excel spreadsheet in the **Supplementary Materials**.

METHODS

Bacteria and PMN Preparation

Staphylococcus aureus strain AH2547 (HG001 + pCM29, courtesy of Alex Horswill), a known biofilm-forming strain (Pabst et al., 2016) with constitutive expression of green fluorescent protein (GFP), was grown overnight in tryptic soy broth supplemented with 10 μg/ml chloramphenicol for maintenance of the GFP carrying pCM29 plasmid. Overnight

cultures were centrifuged for 5 min at 4,000 rpm, rinsed, resuspended in phosphate buffered saline (PBS), and serially diluted. Cells were attached to a 4-chambered glass bottom petri dish (Cellvis, CA, USA) to facilitate live imaging. To attach bacteria, 10 μL of diluted bacterial suspension were added to the surface. After 30 min of incubation at 37°C, unattached bacteria were gently rinsed from the surface with PBS. Each chamber of the petri dish was filled with 1 ml of 10% fresh human serum in Hank's Balanced Salt Solution (HBSS) with Ca^{2+} and Mg^{2+} and incubated at 37°C for 30 min. PMNs were isolated from multiple heparinized venous blood obtained from healthy donors following a standard IRB-approved protocol as described previously (Voyich et al., 2005, 2009). All donors provided written informed consent to participate in the study. PMNs were isolated under endotoxin-free conditions (<25 pg ml^{-1}) and purity (<1% PBMC contamination) and viability (<2% propidium iodide positivity) of PMN preparations were assessed by flow cytometry as previously described (Voyich et al., 2005, 2009). PMNs were kept on ice until stained with LysoBrite Red (AAT Bioquest, CA, USA) according to the manufacturers instructions. Stained PMNs were added to the appropriate wells and interactions between *S. aureus* and PMNs were imaged for 4 h. Some experiments utilized PMNs from the same donor, although PMNs from the same donor were always collected and prepared on different days.

For the antibiotic treatment assay, *S. aureus* was attached to the petri dish surface as described above, grown for 3 h in 10% human serum in HBSS, then challenged with 10 μg/mL gentamicin and imaged for 4 h.

Microscopy

A Leica SP5 inverted confocal laser scanning microscope was utilized for all imaging. GFP-tagged bacteria and PMNs were excited with the 488 nm and 561 nm laser lines, respectively. A LiveCell (Pathology Devices, CA, USA) environmental chamber system was utilized to maintain 5% CO_2, 20% O_2, 50% humidity, and 37°C for sample incubation during imaging. Image stacks 12–20 μm in size with 1-μm z-slices were taken sequentially at 2–3 min intervals over a 4-h time course using a Leica 20x/0.7 NA dry objective lens. At least two fields of view from each chamber of the dish were generally imaged per experiment. FOVs with large numbers of bacterial objects were selected for imaging to visualize the highest number of *S. aureus*—PMN interactions.

Image Analysis

Images were analyzed as maximum projections of z-stacks at each time point. MetaMorph v 7.8.13 (Molecular Devices) image analysis software was used to measure change in bacteria biomass by quantifying the thresholded area of bacterial green fluorescence in each collected image. Log reductions (LR) in GFP area for bacteria treated with PMNs were calculated for each time point using the formula

$$LR = \log_{10}\left(Control\right) - \log_{10}(Treated)$$

where the control area was an average area of the two fields of view from the control well for a given experiment and the treated area was of a single field of view of interest.

Repeatability Analysis

A linear mixed effects model (LMM) was fit to each hour of data [either $\log_{10}$(GFP) area of the controls or LRs] separately, with FOV nested in Experiment as random effects and time as a covariate. This LMM estimated the components of variance due to experiment (Var_{exp}), FOV (Var_{FOV}) and time (Var_{time}). To pool the time periods in each approximate phase of growth, an LMM was fit to data from each phase with fixed effects for log(Inoculum), time period and the two-way interaction, and covariates for time as well as all 2 and 3 way interactions with the fixed effects. The serial correlation over time was modeled with an autoregressive process of order 1 [AR(1)] by fitting the LMM and then inflating the variance for the error term using the approach described in section 15.2 of Ramsey and Schafer (2013). This variance inflation factor is $c = \frac{1+r}{1-r}$ that depends on the first *serial correlation coefficient* $r = \frac{\sum_{t=2}^{n} res_t \times res_{t-1}}{\sum_{t=1}^{n} res_t \times res_t}$ where *res* is the set of LMM residuals. Quadratic trend was fit to the variances as a function of the LR as described by Parker et al. (2018b). All LMMs were fit in R v4.0.1 package lme4 (Bates et al., 2015).

Experimental Designs

The main question of interest is how to use the repeatability results from the LMM to answer questions about microscopy experimental design. To answer this question, we focus on the precision of the mean response [either $\log_{10}$(GFP area) or LR] given by the standard error of the mean (SEM) and a one-sided 95% CI of the mean. The equation for the SEM from an experimental design with n_{exp} experiments, n_{FOV} FOVs per experiment and n_{time} per FOV is

$$SEM = \sqrt{\frac{Var_{exp}}{n_{exp}} + \frac{Var_{FOV}}{n_{exp}n_{FOV}} + \frac{Var_{time} \times c}{n_{exp}n_{FOV}n_{time}}} \quad (1)$$

Equation (1) is the standard formula for the SEM when the data are balanced and there is no serial correlation over time (i.e., when $r = 0$ and the variance inflation factor is $c = 1$). The variance inflation factor is $c > 1$ when the data over time are correlated in which case the SEM is inflated to account for that correlation. In other words, as the variance inflation factor c increases due to higher correlation over time, the SEM also increases through the third term in equation (1). We estimated the variance components Var_{exp}, Var_{FOV}, and Var_{time} in equation (1) from our data using the LMM described above. Our data can be used to estimate the SEM when the images are collected at a single time point (i.e., there is no video) by setting $n_{time} = 1$ in equation (1). If one wants to use their own data for single time point image data, then one would fit an LMM with a single random effect for experiment, retrieve variance components for Var_{exp} and Var_{FOV}, and plug in values for n_{exp} and n_{FOV} and set $Var_{time} = 0$ in equation (1).

For any growth phase given in **Table 1**, we can fill in the values in equation (1) to get the SEM for the mean log(GFP area) for the untreated control biofilms. For example, for an experiment that is run to the early exponential phase (see row 2 in **Table 1**), the

TABLE 1 | Repeatability standard deviation for each phase of biofilm growth and the proportion (Prop.) of variance attributable to each of experiment, field of view, and time.

Phase	Time Period (h)	Prop. Exp	Prop. FOV	Prop. Time	Repeat. SD	First serial correlation	Inflation factor
Lag	[0, 2]	0.3207	0.6707	0.0086	0.4867	0.8313	10.9
Early	[2, 4]	0.5831	0.4033	0.0137	0.3186	0.8751	15.0
Late	[4, 8]	0.4540	0.5151	0.0309	0.2148	0.9129	22.0

The closer that the serial correlation r is to 1, the more dependent the data are over time. The variance inflation factor indicates how much the serial correlation over time increases the SEM via equation (1).

SEM is

$$SEM = \sqrt{\frac{0.583 \times 0.3186^2}{n_{exp}} + \frac{0.403 \times 0.3186^2}{n_{exp} n_{FOV}} + \frac{0.0137 \times 0.3186^2 \times 15}{n_{exp} n_{FOV} n_{time}}}.$$

To fill out the previous formula:

- If an image is collected every 3 min, then 20 images are collected in each FOV over an hour-long experiment. So the number of images inputted into the formula is $n_{time} = 20$.
- If 3 FOVs are collected per experiment, then $n_{FOV} = 3$.
- If there are 2 independent experiments then $n_{exp} = 2$.

Given this setup, then the SEM in this example is

$$SEM = \sqrt{\frac{0.583 \times 0.3186^2}{2} + \frac{0.403 \times 0.3186^2}{2 \times 3} + \frac{0.0137 \times 0.3186^2 \times 15}{2 \times 3 \times 20}} = 0.191.$$

In this example the correlation over time inflates the time component of the SEM by the variance inflation factor $c = 15$ (**Table 1**). For these control data collected from the early exponential phase the variance component due to time is small ($\leq$ 1.4%, see **Table 1**), as is commonly the case, so the SEM remains about the same size whether or not the correlation due to time is accounted for [i.e., whether, in equation (1), $c > 1$ as in **Table 1** or $c = 1$].

The lower limit for an upper one-sided 95% confidence interval (CI) for either the mean control response or the mean LR is calculated by

$$Mean - t \times SEM.$$

The multiplier t is the 95th percentile from a t distribution with $n_{exp} - 1$ degrees of freedom. In our results, we consider $n_{exp} = 1, 2, 3$, and 6 and provide the margin of error, MOE $= t \times SEM$, for a one-sided 95% confidence interval for each of these scenarios coupled with specifications for $n_{time} = 10$ or 40, and $n_{FOV} = 1, 2$, and 6. When $n_{exp} = 1$, we use a t distribution with 0.5 degrees of freedom to give an approximate MOE for a 1-sided 95% CI for the mean. For all SEM and MOE calculations for the mean LR, the inflation factor $c = 12$ was used. This value is the maximum possible value observed for our LR data (see **Supplementary Material**).

The SEM in equation (1) is closely related to the repeatability SD that quantifies the variability of the data across time points, FOV and experiments, by

$$\text{Repeatability SD} = \sqrt{Var_{exp} + Var_{FOV} + Var_{time} \times c} \quad (2)$$

Equation (2) shows that when the data over time are correlated, then the repeatability is inflated to account for that correlation. In other words, as the variance inflation factor c increases due to higher correlation over time, the repeatability SD also increases through the third term in equation (2).

RESULTS

Experimental Setup and Data Acquisition

Previous and ongoing studies in our lab have utilized a time lapse imaging system to analyze interactions between human PMNs and surface adherent *S. aureus* (Ghimire et al., 2019; Pettygrove et al., 2021). Data were mined from two types of experiments that were performed to investigate the effects of PMN density and *S. aureus* aggregate size on the overall clearance of *S. aureus* from a glass surface. We will refer to these two experiment setups as "density" and "head start" experiments, respectively (**Figure 1**). Extensive time-lapse microscopy data of bacterial growth spanning an 8-h time frame and several corresponding treatment conditions were collected for each of the described experiments (**Figures 2**, **3**). The overarching results and interpretation of these experiments will not be discussed at length here but are the subject of a separate manuscript (Pettygrove et al., 2021). For density experiments it was observed that a high concentration of PMNs was required to clear adherent *S. aureus* with the "medium" and "high" treatment conditions resulting in significant clearance of bacteria (**Figures 3A–C**). Head start experiments demonstrated that aggregated *S. aureus* becomes highly resilient to PMN clearance and bacteria will persist through PMN challenge (**Figure 3D**). The relative contributions of inter-experimental, inter-FOV, and inter-timepoint variances at distinct phases in the bacterial growth cycle (namely during lag phase, early exponential phase, and late exponential phase) were all calculated from the collected time lapse microscopy data.

Analysis of Untreated Biofilm

Variability of Early *S. aureus* Biofilm Colonization

Variability in observed GFP area, a surrogate for viable biomass (Schwartz et al., 2009), was analyzed over time for experiments at two inoculation concentrations for 4 h time periods (**Figure 2**).

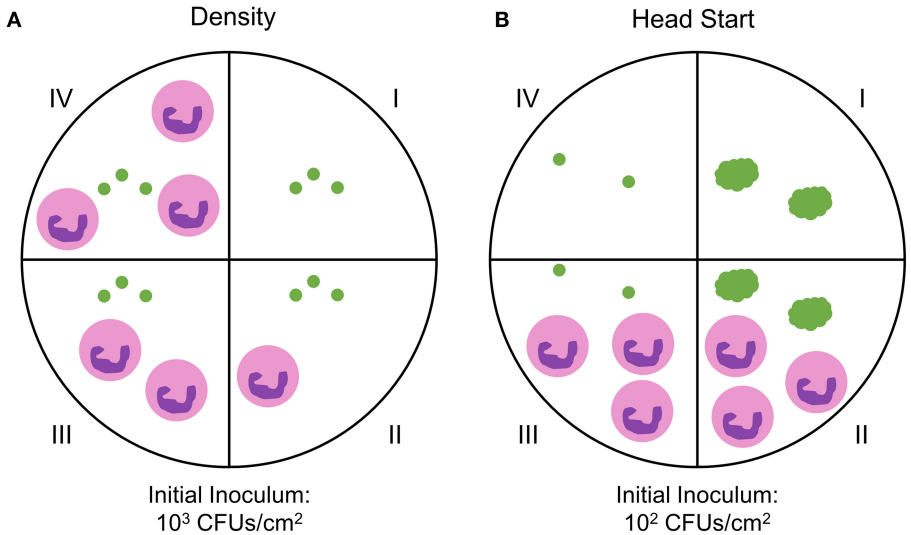


FIGURE 1 | Experimental setup for probing PMN-*S. aureus* interactions. Two FOVs were imaged in each well unless otherwise noted. **(A)** *S. aureus* cells were attached to the dish surface at an initial inoculum of ~10^3 CFUs/cm^2. A single well-contained *S. aureus* only (green dots, well I) while all other wells received *S. aureus* and PMNs (pink circles, wells II - IV). *S. aureus*—PMN interactions were then imaged for 4 h. $N = 10$ independent experiments producing $n = 20$ fields of view (FOV) (control wells), n = 8 FOVs ("low" PMN concentration), $n = 18$ FOVs ("medium" PMN concentration), n = 14 FOVs ("high" PMN concentration). The data are unbalanced for the different neutrophil concentrations. **(B)** *S. aureus* cells were attached to the dish surface at an initial inoculum of ~10^2 CFUs/cm^2. To aggregate cells, *S. aureus* was incubated for 4 h after initial attachment (wells I and II). Single *S. aureus* cells were attached in control wells (wells III and IV) at the 4 h time point for comparison to non-aggregated cells. PMNs were then added to treatment wells (wells II and III) and imaged for 4 h. One well per condition (wells I and IV) was not treated with PMNs. $N = 6$ independent experiments with 12 FOVs per condition.

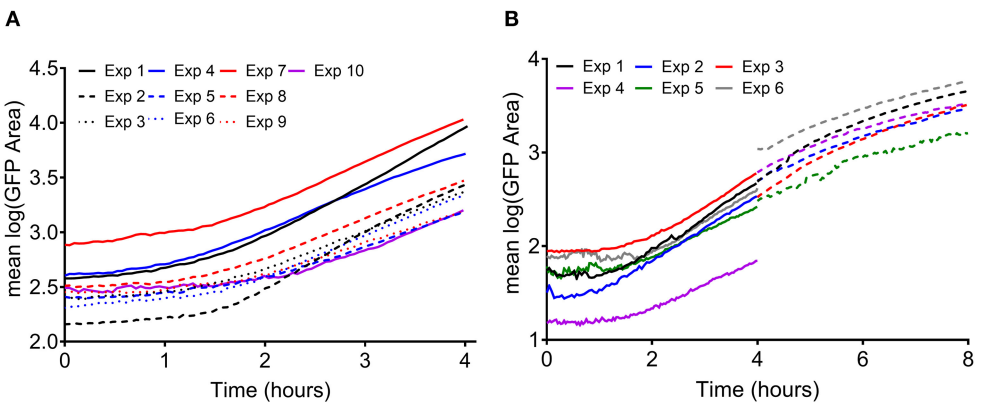


FIGURE 2 | Measured growth of surface adherent *S. aureus* without an antimicrobial treatment. Each line represents the average of two FOVs from a given experiment. **(A)** Bacterial growth measured in density experiments (as in **Figure 1A**, well I). The initial inoculum was measured to be 1.83 ± 1.59 × 10^3 CFUs/cm^2. **(B)** Bacterial growth measured in head start experiments (as in **Figure 1B**). Solid lines indicate control wells in which single cells were attached and observed for 4 h (well IV in **Figure 1B**). Dashed lines indicate wells in which bacteria were given 4 h to grow prior to the start of imaging, then observed for 4 h (hours 4–8, well I in **Figure 1B**). The initial inoculum was measured to be 1.93 ± 0.70 × 10^2 CFUs/cm^2.

Bacterial growth was consistent between experiments, containing a lag phase, early exponential phase, and in the case of the "head start" experiments (**Figure 2B**), a late exponential phase. The repeatability standard deviation (SD) was calculated (data from **Figure 2**) for each hourly period from 0 to 8 h (hours 4–8 from experiments with a head start only), broken out by the inoculation level (**Supplementary Table S1**). Although the biomass area increases dramatically over time as the biofilm aggregates grow (**Figure 2**), the repeatability SDs and the proportions of variance were similar for the two inoculum levels within the lag or early exponential phase. The observed variance was also similar for each of the hourly periods within each phase, regardless of inoculation concentration (i.e., in **Figure 2**, the spread in the curves for the experiments is about the same over time). Thus, the data were pooled across inoculum levels and hours for each phase of growth. A single repeatability SD

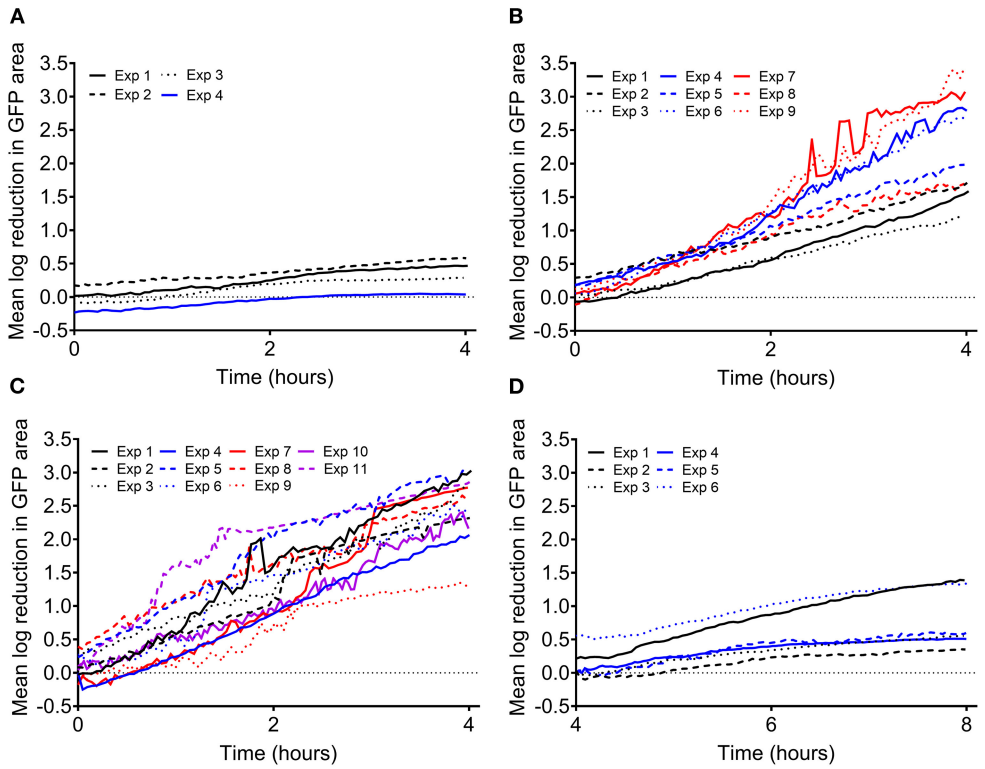


FIGURE 3 | Log reduction in surface attached *S. aureus* biomass during PMN challenge compared to the average of the two corresponding control FOVs (no PMNs). Each line represents the average of two FOVs from a given experiment unless otherwise noted. Data shown from **(A)** density experiments with a low number of PMNs ($2.39 \pm 1.13 \times 10^3$ PMNs/cm^2, 8 FOVs, depicted in **Figure 1A**, well II) and **(B)** density experiments with a medium number of PMNs ($1.53 \pm 1.25 \times 10^4$ PMNs/cm^2, 18 FOVs, depicted in **Figure 1A**, well III). **(C)** Data shown from density experiments with high PMNs (Exps 1–5, $2.57 \pm 1.18 \times 10^4$ PMNs/cm^2, 14 FOVs, depicted in **Figure 1A**, well IV) and control wells from head start experiments containing single cells and PMNs (Exps 6–11, $1.09 \pm 0.31 \times 10^4$ PMNs/cm^2, 12 FOVs, depicted in **Figure 1B**, well III). Except for experiment 5, which contains 6 FOVs of the same condition, each line represents the average of two FOVs from a given experiment. **(D)** Data shown from head start experiments where aggregated bacteria were challenged with PMNs ($1.72 \pm 0.93 \times 10^4$ PMNs/cm^2, 12 FOVs, depicted in **Figure 1B**, well II).

and components of variance of the log(area) were calculated for each phase of growth (**Table 1**). Overall, variation decreased as cells progressed through exponential phase. Thus, seeding and culture conditions may affect variation more during the early periods of biofilm colonization. It is crucial to have side-by-side controls in all experiments that are highly repeatable, otherwise the log reductions can be affected (Stewart, 2015). The level of variability of the untreated controls from our experiments is acceptably small (a repeatability SD of untreated controls <½ log is acceptable) compared to other methods in the peer reviewed literature that have been standardized with AOAC International (Tilt and Hamilton, 1999) or ASTM International (Parker et al., 2014; Goeres et al., 2019, 2020; Allkja et al., 2021).

Design Recommendations

After calculating the variation in *S. aureus* biomass during different phases of growth (**Table 1**), the improvement in the standard error of the mean $\log_{10}$(GFP Area) (SEM) was estimated under different experimental designs using the formula for SEM outlined in equation (1) in the methods. Given the SEM, we then predicted the margin of error (MOE) for a 95% confidence interval for the mean log(GFP Area) of untreated biofilms under varied numbers of replicate experiments (1–6), FOVs (1, 2, 6) and time rates of acquisition (10, 40 images per hour) (**Table 2**). Larger SEM and MOE values indicate more imprecision when estimating biomass from an image. Unsurprisingly, results for early exponential phase data indicate that at least three experiments should be performed to reach a MOE below 1 log(GFP area). While one FOV may be acceptable, exceeding two FOVs per experiment yields only mild returns with regards to increasing precision, suggesting two FOVs may be optimal. Additionally, increasing the temporal resolution only negligibly increases precision, so utilizing fewer timepoints per hour for temporal studies may be advantageous to reduce fluorophore photosensitivity or phototoxicity, especially when exciting with ultraviolet light. Assessments of different experimental designs for lag and late exponential phase growth data produced similar conclusions (**Table 2**). Overall, the improvement in precision diminishes with increased numbers of experiments or FOVs in each experiment. However, there is a much greater decrease in MOE when more experiments are performed, even when

TABLE 2 | Expected margin of error for a one-sided 95% confidence interval for the mean $\log_{10}$(GFP area) at different phases of growth and a variety of experimental parameters for untreated controls.

Lag	**FOV**	**1**	**1**	**2**	**2**	**6**	**6**
	Frames	**10**	**40**	**10**	**40**	**10**	**40**
Experiment number	1	20.0167	19.9466	16.2652	16.2221	13.1839	13.1661
	2	2.1765	2.1689	1.7686	1.7639	1.4335	1.4316
	3	0.8211	0.8182	0.6672	0.6654	0.5408	0.5401
	4	0.5723	0.5703	0.465	0.4638	0.3769	0.3764
	5	0.4639	0.4623	0.377	0.376	0.3056	0.3051
	6	0.4016	0.4002	0.3264	0.3255	0.2645	0.2642
Early exponential	**FOV**	**1**	**1**	**2**	**2**	**6**	**6**
	Frames	**10**	**40**	**10**	**40**	**10**	**40**
Experiment number	1	13.1355	13.0347	11.6722	11.6156	10.5849	10.5641
	2	1.4283	1.4173	1.2692	1.263	1.1509	1.1487
	3	0.5388	0.5347	0.4788	0.4765	0.4342	0.4333
	4	0.3755	0.3726	0.3337	0.3321	0.3026	0.302
	5	0.3044	0.3021	0.2705	0.2692	0.2453	0.2448
	6	0.2636	0.2615	0.2342	0.2331	0.2124	0.212
Late exponential	**FOV**	**1**	**1**	**2**	**2**	**6**	**6**
	Frames	**10**	**40**	**10**	**40**	**10**	**40**
Experiment number	1	8.9894	8.7664	7.6221	7.4911	6.554	6.5035
	2	0.9774	0.9532	0.8288	0.8145	0.7126	0.7071
	3	0.3687	0.3596	0.3126	0.3073	0.2688	0.2668
	4	0.257	0.2506	0.2179	0.2142	0.1874	0.1859
	5	0.2083	0.2032	0.1767	0.1736	0.1519	0.1507
	6	0.1804	0.1759	0.1529	0.1503	0.1315	0.1305

The table shows that when quantifying untreated biofilms with a mean $\log_{10}$(GFP Area) during lag phase, at least 4 experiments with 2 FOVs each must be performed to get the error bars less than one half log. For early exponential phase, at least 3 experiments with 2 FOVs in each or 4 experiments must be performed while at least 3 experiments must be performed for late exponential phase.

fewer FOVs are imaged in each experiment compared to fewer experiments with a large number of FOVs in each. For example, when estimating the mean $\log_{10}$(GFP area) of the untreated control biofilms during the early exponential phase at 95% confidence, consider a design utilizing two replicate experiments, two FOVs, and ten images per hour. Tripling the number of FOVs to six only yields a 9.3% reduction in the margin of error, while performing an additional experiment with two FOVs in each experiment results in a 62.3% reduction in the margin of error (**Table 2**). Conversely, performing six experiments with two FOVs and 10 images per hour reduces the margin of error by 81.5%, but requires a 3-fold increase in effort. Thus, our results suggest minimal benefit to utilizing a large number of FOVs with fewer experiments and instead recommend increasing the number of experiments performed when observing untreated biofilms during the early stages of colonization and development. However, the improvement in precision from performing more experiments must be weighed against the time and resource commitment required for those repetitions.

Analysis of Treated Biofilm

Variability of PMN Clearance of Nascent *S. aureus* Biofilm

By analyzing the results of experiments where surface attached *S. aureus* cells were challenged by human PMNs, we can make experimental design recommendations for studying the effect of an antimicrobial treatment on the viability of early biofilms. The two experimental setups (**Figure 1**) allowed us to analyze two different treatment scenarios. Experiments where the PMN density was varied (**Figure 1A**) can serve as an analog to changing the concentration of an antimicrobial treatment while the "head start" experiments (**Figure 1B**) can be modeled as application of a treatment at two different stages in biofilm development (in this case, lag and late exponential phase). It is important to note that discovery and subsequent killing of sparsely seeded *S. aureus* cells on a surface is a stochastic process that is highly dependent on the PMN concentration, unlike an antibiotic treatment that is comparatively homogenously applied (Ghimire et al., 2019). We calculated LRs in GFP area within each treatment FOV compared to the average of the two corresponding control FOVs at each

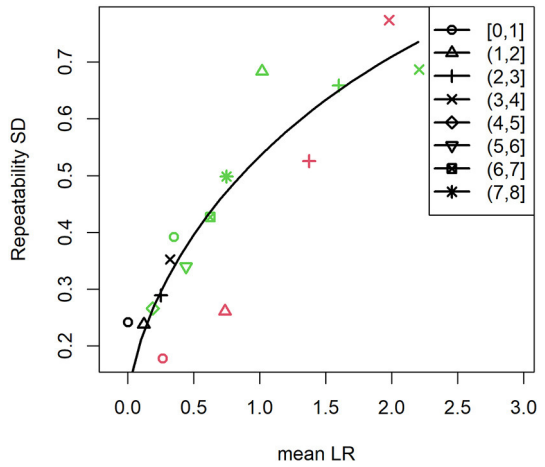


FIGURE 4 | Repeatability standard deviation increases with the observed log reduction. Symbol shapes correspond to the observation time period (h) while colors indicate the concentration of PMNs that were used to challenge *S. aureus*. Black symbols indicate low PMN concentration (**Figure 3A**), red symbols indicate medium PMN concentration (**Figure 3B**), and green symbols indicate high PMN concentration (**Figures 3C,D**). Solid line represents the quadratic trend in the repeatability variance. Brackets indicate that the value is included in the given time range while parentheses indicate that the value is excluded.

time point. For PMN challenges against single cell *S. aureus*, as PMN density was increased, we observed higher LRs in GFP area, but increasing variability as the magnitude of the LR increased (**Figures 3A–C**). We similarly observed higher variability in LR in the efficacious "high PMN treatment" against single cells as compared to the relatively ineffective "high treatment" against aggregated cells (**Figure 3D**). Repeatability SD did not change markedly over time, but did increase with the observed LR, regardless of the experimental setup (Density vs. Head Start) (**Figure 4**). Intuitively, a treatment that is poorly efficacious (**Figures 3A,D**) will consistently result in small reductions and be more repeatable (Parker et al., 2018a). On the other hand, for a treatment that is highly efficacious and consistently kills all attached cells, we would expect a high level of repeatability as well (such a highly efficacious treatment was not tested in our experiments). Here, we observe that treatments of middling efficacy are highly variable. Thus, experiments where a treatment is expected to be highly efficacious or very ineffective will require fewer repetitions while moderately effective treatments require more. Our statistical model for the LRs explicitly accounts for differences among experiments. Because multiple donors contributed PMNs, the repeatability variance in **Figure 4** may be larger than what would be expected if all PMNs were collected from the same donor. Nonetheless, the repeatability SDs for LRs that were observed in our system were <1 log and hence acceptable according to ASTM International and AOAC international (Parker et al., 2018a).

Design Recommendations

In our experiments measuring killing of attached *S. aureus* cells by human PMNs, we observed that the variance changes as a function of the LR regardless of inoculation level and time (**Figure 4** and **Supplementary Table S2**). As with the data from control wells, we used these results to evaluate how frequently to image, how many experiments to conduct, and how many FOVs to utilize per experiment in order to generate estimates of the mean LR with statistical confidence. The first step was to calculate the SEM for the mean LR using equation (1) from the estimated variance components for experiment, FOV and time (**Supplementary Figure S1**). Using these values, the expected MOE was assessed for experimental designs for treatments that have a mean LR = 0.5, 1, or 2 over a wide range of experiment numbers, FOV numbers, and temporal resolutions (**Table 3**). Because the variance increases as the mean LR increases (**Figure 4**), the SEM increases as the LR increases. Nevertheless, this does not translate into more experimental effort to precisely estimate the mean LR for moderately efficacious treatments (with mean LR = 2) compared to less efficacious treatments (with mean LRs $\leq$ 1) because the magnitude of the LR is less for less efficacious treatments. For an expected LR = 0.5, at least 4 experiments are required to detect a significant kill, while for LR = 1 or 2, only 3 experiments are necessary (**Table 3**). While the minimum number of experiments required to detect a significant kill does not change based on the number of FOVs (except when using 6 FOVs when LR = 0.5), modest improvements in the MOE are expected when at least 2 FOVs are utilized and may provide considerable benefit to the researcher. Providing an expected LR is a standard input for sample size calculations. The only difference in our approach from the standard approach is that in addition to using the expected LR as the margin of error input into the sample size calculator, we also use the expected LR to look up the variability of the LR (*via* **Figure 4**). Using the standard sample size calculator approach, the variability would be assumed to be the same for all LR values, which clearly is not the case for these efficacy data (see **Figure 4** and the review in Parker et al., 2018a).

Validation of Design Recommendations

To determine how well **Table 3** predicts the margin of error, we assessed a second dataset utilizing a 10 μg/mL gentamicin treatment rather than PMN treatment. Similar to the previously described experiments, *S. aureus* was grown for 3 h, then treated with gentamicin for 4 h and imaged via confocal microscopy. We expected a LR of around 1 after 4 h of contact time and a repeatability SD of around 0.5 (**Figure 4**). **Table 3** suggests that 3 independent experiments with 1 FOV each would be sufficient to detect a significant kill. The dataset that we analyzed contained 3 independent experiments with 3 control FOVs and 3 treated FOVs in each experiment, allowing us to assess the MOE for many possible combinations of 3 experiments with 1 control FOV and 1 treated FOV (27 possible control FOV combinations and 27 possible treatment FOV combinations yielding 729 possible LRs). We simulated 1,000 studies with this design via bootstrap sampling and calculated the mean LR, MOE, and lower 95% confidence limit for each experimental design (**Supplementary Figure S2**). Overall, the gentamicin treated data was less variable (mean LR = 0.80, repeatability SD = 0.12) than the neutrophil treated data (repeatability close to 0.5 according to **Figure 4**), resulting in MOEs much smaller than the predicted value of 0.91 (**Table 3**).

TABLE 3 | Expected margin of error (MOE) for a 1-sided 95% confidence interval for the mean log reduction in GFP area at a variety of experimental parameters and when LR = 0.5 (i.e., when percent reduction is 68%), when LR = 1 (i.e., when percent reduction is 90%), and when LR = 2 (i.e., when percent reduction is 99%).

LR = 0.5	**FOV**	**1**	**1**	**2**	**2**	**6**	**6**
	Frames	**10**	**40**	**10**	**40**	**10**	**40**
Experiment number	1	16.2871	16.2155	13.872	13.83	11.9948	11.9786
	2	1.7681	1.7604	1.506	1.5014	1.3022	1.3004
	3	0.6681	0.6651	0.569	0.5673	0.492	0.4913
	4	0.4656	0.4636	0.3966	0.3954	0.3429	0.3425
	5	0.3775	0.3758	0.3215	0.3205	0.278	0.2776
	6	0.3268	0.3254	0.2783	0.2775	0.2407	0.2403
LR = 1	**FOV**	**1**	**1**	**2**	**2**	**6**	**6**
	Frames	**10**	**40**	**10**	**40**	**10**	**40**
Experiment number	1	22.1521	22.1521	18.9751	18.9751	16.5211	16.5211
	2	2.4048	2.4048	2.06	2.06	1.7935	1.7935
	3	0.9086	0.9086	0.7783	0.7783	0.6777	0.6777
	4	0.6333	0.6333	0.5425	0.5425	0.4723	0.4723
	5	0.5134	0.5134	0.4398	0.4398	0.3829	0.3829
	6	0.4445	0.4445	0.3807	0.3807	0.3315	0.3315
LR = 2	**FOV**	**1**	**1**	**2**	**2**	**6**	**6**
	Frames	**10**	**40**	**10**	**40**	**10**	**40**
Experiment number	1	29.3802	28.4735	24.0585	23.5071	19.7287	19.5059
	2	3.1895	3.0911	2.6118	2.5519	2.1418	2.1176
	3	1.2051	1.1679	0.9868	0.9642	0.8092	0.8001
	4	0.8399	0.814	0.6878	0.672	0.564	0.5577
	5	0.6809	0.6599	0.5576	0.5448	0.4572	0.4521
	6	0.5895	0.5713	0.4827	0.4717	0.3959	0.3914

The MOE was calculated from equation (1) using the variance components from ***Supplementary Figure 1*** *for each LR value. The table shows that when studying a treatment with a mean LR = 0.5, at least 3 experiments, each with 6 fields of view; or 4 experiments with any number of fields of view must be performed to be able to detect a statistically significant kill. When studying a treatment with a mean LR = 1 or LR = 2, at least 3 experiments must be performed to be able to detect a statistically significant kill. Because the variance due to time is 0 for LR = 1 (****Supplementary Figure 1****), there is no effect on the margin of error due to differences in the number of images acquired.*

In 100% of the simulations, the design produced a statistically significant kill (**Supplementary Figure S2C**). This indicates that the recommendation of 3 experiments and 1 FOV, while conservative, was sufficient for detecting the mean LR with 95% confidence using this experimental setup (gentamicin treatment rather than PMN treatment).

DISCUSSION

Real-time CLSM non-invasively provides quantitative insight into the dynamics of living biofilms (Stewart et al., 2009; Davison et al., 2010; McLoon et al., 2011; Hung et al., 2013; Franklin et al., 2015; Ghimire et al., 2019; Pettygrove et al., 2021). Continued exploration in this area by our lab has yielded an abundance of time-lapse microscopy data that can inform the proper design of microscopy studies (Ghimire et al., 2019; Pettygrove et al., 2021). The data included are particularly suitable for assisting with the experimental design of studies investigating initial bacterial attachment, early biofilm formation, and biocidal treatments of bacteria in this stage. However, the development of criteria for selecting the appropriate number of FOV that are required to produce representative quantitative results have been overlooked for this type of data collection. In this paper, we aimed to retrospectively utilize time-lapse data to provide quantitative guidance on the aspects of experimental design such as the numbers of experiments, fields of view, and images over time to provide an acceptable level of precision when estimating biomass on a surface.

In the presented data, variance for untreated controls representative of early biofilm formation (i.e., early colonization), depended on the growth phase of interest (lag or exponential in our data). The repeatability SD tended to decrease as cells entered exponential phase (**Table 1**). Thus, when designing an imaging experiment to study biofilm colonization, the user should primarily consider which phase of growth will be monitored and adapt the number of experiments, FOVs, and time points to include in the experimental design accordingly. The increased repeatability SD observed during the lag phase compared to early and late exponential is likely in part due to the high amounts of heterogeneity in the initial observed growth. *S. aureus* cells were

seeded onto the surface in PBS, rather than a nutrient rich media, resulting in differential recovery from lag phase for individual colonies, leading to high intra-FOV variation during lag phase as attached cells began to readjust to the media conditions. Thus, initial seeding conditions may affect variability as well and must be assessed for each system. For experiments where the measured output is a reduction in viable bacteria (GFP area in these experiments), the expected LR is required to estimate the appropriate number of experiments, FOVs, and time points. It is to be expected that the variability will be markedly different for different laboratory systems (Parker et al., 2018a; Stewart and Parker, 2019). Nonetheless, here we present a methodology for others to follow to estimate the variance in their systems and allow experimental design assessments.

As it pertains to repeatability, experimental bias is a significant obstacle for imaging studies. Sources of bias in imaging experiments can include the selection of representative regions of interest, image collection parameters, and the number of visualized FOVs (Lee and Kitaoka, 2018). Relatively arbitrary recommendations have been made for approaches to minimize bias such as visualizing a greater number of FOV, increasing the size of the visualized area, and imaging multi-well plates (Patel and McGhee, 2013; Lee and Kitaoka, 2018). There is limited suitability of some of these recommendations as it pertains to real-time biofilm imaging due to phototoxicity and the need for rapid image acquisition. As is common, the fields of view that we analyze here were not randomly chosen since the bacterial attachment was not homogenous across the surface. Statistical predictions and confidence intervals from hypothetical experimental designs, as we do here, presume that the data collected are representative of some population of all possible fields of view that could have been selected. Usually, representative data are assured by random sampling, in this case randomly selecting fields of view, wherein the population of interest is all possible fields of view. True random sampling is not typically performed in imaging studies as many fields of view are empty, especially during attachment or early biofilm formation. For the data collected here, the population of interest was all fields of view that contained some population of PMNs and biofilm aggregates, as determined by the microscopist. Fields of view were selected to maximize the number of PMN-*S. aureus* interactions that could be imaged, thus most fields of view were selected because of a high number of *S. aureus* aggregates present. While FOV selection is subjective and often dependent on the microscopist, our results are a starting point for estimating the sources contributing to the variance in observed biomass and designing microscopy experiments. There are methods, however, to reduce bias due to the subjective selection of fields of view. For example, in the ecological literature, there are adaptive sampling techniques that begin by randomly selecting locations for observation, i.e., randomly selecting a FOV (Thompson, 1990). When the ecologist (or microscopist) moves and encounters a new location (or FOV) of interest, the resulting statistics can be "adapted" to account for this deviation in sampling. Implementation of adaptive sampling that adjusts the image statistics is a matter for future research. Our lack of random sampling is a potential weakness of our work. We recommend adaptive sampling of FOVs to minimize bias due to subjective FOV selection, but until the tools for CLSM have been developed, the next best approach to alleviate bias is to randomly select fields of view from the surface being imaged.

Our results, not surprisingly, point to the need to perform at least three experiments in most scenarios. For a given growth phase or expected LR, a substantial contribution to variance in biomass is due to differences between experiments (**Table 1** and **Supplementary Figure S1**). Furthermore, our results indicate that a single FOV in each experiment may be sufficient, as the resulting reduction in variance from increasing FOVs is small. Most experimenters will want to collect at least two FOVs, but our results suggest that there is only minimal benefit to using more than two in most cases. Furthermore, as neutrophil challenge is a relatively heterogeneous antimicrobial treatment as compared to antibiotic therapy, it is plausible that the inter-FOV variance in our system is higher than might be observed in an antibiotic treated system. Lastly, our results demonstrate that there is not substantial increase to precision by increasing the temporal resolution of the acquired images in the video. In our variance predictions, little difference was seen between conditions utilizing 10 images per hour compared to 40 images per hour, suggesting an experimenter should utilize the temporal resolution that is best for their sample as the effect on their data quality will be minimal, barring phototoxicity. For visualizing rapid processes, high temporal resolution has an obvious benefit, however lower temporal resolution can both decrease phototoxicity and downstream image analysis burden. We observed little phototoxicity or fluorophore bleaching in our system despite a frequent image capture rate, however each experimenter should decide how rapidly to image based on the organism, fluorophores, and conditions necessary for their work. It is to be expected that a study of a different organism, even when using the same method, will produce results that have different levels and sources of variation than *S. aureus*. For example, a motile organism may enter and leave a FOV, while other organisms may have less predictable attachment mechanics. Similarly, different growth or treatment conditions would introduce different levels and sources of variability. These factors could lead to different inter-experimental, inter-FOV, or inter-timepoint variance compared to the values measured in our study. Therefore, we strongly recommend that researchers generate their own pilot data using their own methods and organisms and use the resulting variance component estimates to complete equation (1) via the included Excel tool to generate study designs. However, in the absence of available pilot data, which unfortunately is many times the case, the study designs (in **Tables 2**, **3**) calculated from our data can serve as a starting point, with the understanding that there are likely differences between the system we present here and systems used by other researchers. Regardless of the data used, the analysis framework presented in this manuscript can help other researchers design statistically sound experiments based on a quantitative understanding of the contributors to variance in their system. Put another way, while the laboratory methodology may vary, the same statistical approach can be utilized to assess study designs.

The **Supplementary Materials** contains an Excel tool that allows the user to predict the number of experiments, FOVs, and images per hour (i.e., an experimental design) to estimate experimental parameters calculated from images of early biofilms at 95% confidence. The two parameters we focus on are the log transformed surface coverage [i.e., $\log_{10}$(GFP area)] of untreated control biofilms or the LR measure of treatment efficacy. The tool requires estimates of the variance components due to experiments, FOVs and time for different experimental setups. For untreated control biofilms, the variance estimation will depend on the growth phase of interest. For treated biofilms, the variance estimation will depend on the expected LR.

As the field of quantitative imaging progresses there is a growing need for increased reporting on imaging acquisition parameters, processing, figure generation, and statistical analysis. The results reported here provide a statistical approach for deciding on appropriate replication to perform in imaging-based experiments during early biofilm colonization. Results from the presented approach also provide an efficient way for communicating methodology and statistical rigor from imaging experiments to the scientific community. While the variance observed will depend on one's specific experimental setup, these results do demonstrate the value of prioritizing experimental repetitions over high numbers of FOVs. The results (**Tables 2**, **3**) also demonstrate the diminishing returns gained from increasing experimental repetitions and FOVs, which can help researchers decide how best to balance workload and the expected quality of data. The data mined for these analyses is derived from experiments examining sparsely seeded surfaces, thus future experiments should examine densely seeded surfaces and other measurable outputs such as volume rather than just area or surface coverage. In addition to early colonization studies there is a large component of biofilm research that is conducted on mature biofilms, hence future work should investigate the variability observed in imaging mature biofilms. Nonetheless, the statistical approach presented herein could be applied to biofilms at different life stages or different organisms and assays. We intend for this method to serve as a framework to be utilized with any pilot data examining biofilm growth to better inform the experimental design. Together, the results described in this paper can act as a roadmap to assist researchers in the design of microscopy-based studies of biofilm colonization to ensure that acquired data is representative of the sample and statistically robust.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Montana State University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

BP designed and conducted experiments, performed data analysis, and drafted the manuscript. HS conceived the study, performed data analysis, drafted the manuscript, and supervised the study. KP performed blood draws and PMN isolations and revised the manuscript. JV and PS designed experiments, provided laboratory resources, revised the manuscript, supervised the study, and acquired funding. AP conceived the study, performed the statistical analysis, drafted the manuscript, and supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants U54GM115371, RO1AI149491, and 1R56AI155692 from the National Institutes of Health. Imaging was made possible by microscope facilities at the Center for Biofilm Engineering, which were supported by funding from the National Science Foundation MRI Program and the M. J. Murdock Charitable Trust.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.785182/full#supplementary-material

REFERENCES

Allesen-Holm, M., Barken, K. B., Yang, L., Klausen, M., Webb, J. S., Kjelleberg, S., et al. (2006). A characterization of DNA release in *Pseudomonas aeruginosa* cultures and biofilms. *Mol. Microbiol.* 59, 1114–1128. doi: 10.1111/j.1365-2958.2005.05008.x

Allkja, J., van Charante, F., Aizawa, J., Reigada, I., Guarch-Pérez, C., Vazquez-Rodriguez, J. A., et al. (2021). Interlaboratory study for the evaluation of three microtiter plate-based biofilm quantification methods. *Sci. Rep.* 11:13779. doi: 10.1038/s41598-021-93115-w

Almstrand, R., Daims, H., Persson, F., Sörensson, F., and Hermansson, M. (2013). New methods for analysis of spatial distribution and coaggregation of microbial populations in complex biofilms. *Appl. Environ. Microbiol.* 79, 5978–5987. doi: 10.1128/AEM.01727-13

Baird, F. J., Wadsworth, M. P., and Hill, J. E. (2012). Evaluation and optimization of multiple fluorophore analysis of a *Pseudomonas aeruginosa* biofilm. *J. Microbiol. Methods* 90, 192–196. doi: 10.1016/j.mimet.2012.05.004

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Baudin, M., Cinquin, B., Sclavi, B., Pareau, D., and Lopes, F. (2017). Understanding the fundamental mechanisms of biofilms development and dispersal: BIAM (Biofilm Intensity and Architecture Measurement), a new tool for studying biofilms as a function of their architecture and fluorescence intensity. *J. Microbiol. Methods* 140, 47–57. doi: 10.1016/j.mimet.2017.06.021

Bridier, A., Dubois-Brissonnet, F., Greub, G., Thomas, V., and Briandet, R. (2011). Dynamics of the action of biocides in *Pseudomonas aeruginosa* biofilms. *Antimicrob. Agents Chemother.* 55, 2648–2654. doi: 10.1128/AAC.01760-10

Brileya, K. A., Camilleri, L. B., and Fields, M. W. (2014). "3D-fluorescence *in situ* hybridization of intact, anaerobic biofilm," in *Engineering and Analyzing Multicellular Systems: Methods and Protocols* Methods in Molecular Biology, eds L. Sun and W. Shou (New York, NY: Springer) 189–197. doi: 10.1007/978-1-4939-0554-6_13

Connolly, J. M., Jackson, B., Rothman, A. P., Klapper, I., and Gerlach, R. (2015). Estimation of a biofilm-specific reaction rate: kinetics of bacterial urea hydrolysis in a biofilm. *NPJ Biofilms Microbiomes* 1, 1–8. doi: 10.1038/npjbiofilms.2015.14

Daims, H., Lücker, S., and Wagner, M. (2006). daime, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* 8, 200–213. doi: 10.1111/j.1462-2920.2005.00880.x

Daims, H., and Wagner, M. (2011). "Chapter eight - *in situ* techniques and digital image analysis methods for quantifying spatial localization patterns of nitrifiers and other microorganisms in biofilm and flocs," in *Methods in Enzymology Research on Nitrification and Related Processes, Part B*, eds M. G. Klotz and L. Y. Stein (Academic Press). doi: 10.1016/B978-0-12-386489-5.00008-7

Davison, W. M., Pitts, B., and Stewart, P. S. (2010). Spatial and temporal patterns of biocide action against *Staphylococcus epidermidis* biofilms. *Antimicrob. Agents Chemother.* 54, 2920–2927. doi: 10.1128/AAC.01734-09

de Beer, D., Stoodley, P., Roe, F., and Lewandowski, Z. (1994). Effects of biofilm structures on oxygen distribution and mass transport. *Biotechnol. Bioeng.* 43, 1131–1138. doi: 10.1002/bit.260431118

Franklin, M. J., Chang, C., Akiyama, T., and Bothner, B. (2015). New technologies for studying biofilms. *Microbiol. Spectr.* 3:16. doi: 10.1128/microbiolspec.MB-0016-2014

Ghimire, N., Pettygrove, B. A., Pallister, K. B., Stangeland, J., Stanhope, S., Klapper, I., et al. (2019). Direct microscopic observation of human neutrophil-*Staphylococcus aureus* interaction *in vitro* suggests a potential mechanism for initiation of biofilm infection on an implanted medical device. *Infect. Immun.* 87:19. doi: 10.1128/IAI.00745-19

Gloag, E. S., Turnbull, L., Huang, A., Vallotton, P., Wang, H., Nolan, L. M., et al. (2013). Self-organization of bacterial biofilms is facilitated by extracellular DNA. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11541–11546. doi: 10.1073/pnas.1218898110

Goeres, D. M., Hamilton, M. A., Beck, N. A., Buckingham-Meyer, K., Hilyard, J. D., Loetterle, L. R., et al. (2009). A method for growing a biofilm under low shear at the air–liquid interface using the drip flow biofilm reactor. *Nat. Protoc.* 4, 783–788. doi: 10.1038/nprot.2009.59

Goeres, D. M., Loetterle, L. R., Hamilton, M. A., Murga, R., Kirby, D. W., and Donlan, R. M. (2005). Statistical assessment of a laboratory method for growing biofilms. *Microbiology* 151, 757–762. doi: 10.1099/mic.0.27709-0

Goeres, D. M., Parker, A. E., Walker, D. K., Meier, K., Lorenz, L. A., and Buckingham-Meyer, K. (2020). Drip flow reactor method exhibits excellent reproducibility based on a 10-laboratory collaborative study. *J. Microbiol. Methods* 174:105963. doi: 10.1016/j.mimet.2020.105963

Goeres, D. M., Walker, D. K., Buckingham-Meyer, K., Lorenz, L., Summers, J., Fritz, B., et al. (2019). Development, standardization, and validation of a biofilm efficacy test: The single tube method. *J. Microbiol. Methods* 165:105694. doi: 10.1016/j.mimet.2019.105694

Gomes, I. B., Meireles, A., Gonçalves, A. L., Goeres, D. M., Sjollema, J., Simões, L. C., et al. (2018). Standardized reactors for the study of medical biofilms: a review of the principles and latest modifications. *Crit. Rev. Biotechnol.* 38, 657–670. doi: 10.1080/07388551.2017.1380601

Hall-Stoodley, L., Costerton, J. W., and Stoodley, P. (2004). Bacterial biofilms: from the Natural environment to infectious diseases. *Nat. Rev. Microbiol.* 2, 95–108. doi: 10.1038/nrmicro821

Heydorn, A., Ersbøll, B. K., Hentzer, M., Parsek, M. R., Givskov, M., and Molin, S. (2000). Experimental reproducibility in flow-chamber biofilms. *Microbiology* 146, 2409–2415. doi: 10.1099/00221287-146-10-2409

Hung, C., Zhou, Y., Pinkner, J. S., Dodson, K. W., Crowley, J. R., Heuser, J., et al. (2013). *Escherichia coli* biofilms have an organized and complex extracellular matrix structure. *mBio* 4:645. doi: 10.1128/mBio.00645-13

Korber, D. R., Lawrence, J. R., Hendry, M. J., and Caldwell, D. E. (1993). Analysis of spatial variability within mot+ and mot– pseudomonas fluorescens biofilms using representative elements. *Biofouling* 7, 339–358. doi: 10.1080/08927019309386264

Larimer, C., Winder, E., Jeters, R., Prowant, M., Nettleship, I., Addleman, R. S., et al. (2016). A method for rapid quantitative assessment of biofilms with biomolecular staining and image analysis. *Anal. Bioanal. Chem.* 408, 999–1008. doi: 10.1007/s00216-015-9195-z

Lawrence, J. R., and Neu, T. R. (1999). "[9] Confocal laser scanning microscopy for analysis of microbial biofilms," in *Methods in Enzymology Biofilms* (Academic Press). doi: 10.1016/S0076-6879(99)10011-9

Lee, J.-Y., and Kitaoka, M. (2018). A beginner's guide to rigor and reproducibility in fluorescence imaging experiments. *Mol. Biol. Cell* 29, 1519–1525. doi: 10.1091/mbc.E17-05-0276

Luo, T. L., Eisenberg, M. C., Hayashi, M. A. L., Gonzalez-Cabezas, C., Foxman, B., Marrs, C. F., et al. (2018). A sensitive thresholding method for confocal laser scanning microscope image stacks of microbial biofilms. *Sci. Rep.* 8, 1–14. doi: 10.1038/s41598-018-31012-5

Luo, T. L., Hayashi, M., Zsiska, M., Circello, B., Eisenberg, M., Gonzalez-Cabezas, C., et al. (2019). Introducing BAIT (Biofilm Architecture Inference Tool): a software program to evaluate the architecture of oral multi-species biofilms. *Microbiology* 165, 527–537. doi: 10.1099/mic.0.000761

Magana, M., Sereti, C., Ioannidis, A., Mitchell, C. A., Ball, A. R., Magiorkinis, E., et al. (2018). Options and limitations in clinical investigation of bacterial biofilms. *Clin. Microbiol. Rev.* 31:16. doi: 10.1128/CMR.00084-16

Malone, M., Goeres, D. M., Gosbell, I., Vickery, K., Jensen, S., and Stoodley, P. (2017). Approaches to biofilm-associated infections: the need for standardized and relevant biofilm methods for clinical applications. *Expert Rev. Anti Infect. Ther.* 15, 147–156. doi: 10.1080/14787210.2017.1262257

Martínez-García, R., Nadell, C. D., Hartmann, R., Drescher, K., and Bonachela, J. A. (2018). Cell adhesion and fluid flow jointly initiate genotype spatial distribution in biofilms. *PLOS Comput. Biol.* 14:e1006094. doi: 10.1371/journal.pcbi.1006094

McLoon, A. L., Kolodkin-Gal, I., Rubinstein, S. M., Kolter, R., and Losick, R. (2011). Spatial Regulation of histidine kinases governing biofilm formation in Bacillus subtilis. *J. Bacteriol.* 193, 679–685. doi: 10.1128/JB.01186-10

Menzel, F., Conradi, B., Rodenacker, K., Gorbushina, A. A., and Schwibbert, K. (2016). Flow chamber system for the statistical evaluation of bacterial colonization on materials. *Materials* 9:770. doi: 10.3390/ma9090770

Milferstedt, K., Pons, M.-N., and Morgenroth, E. (2009). Analyzing characteristic length scales in biofilm structures. *Biotechnol. Bioeng.* 102, 368–379. doi: 10.1002/bit.22075

Mueller, L. N., de Brouwer, J. F., Almeida, J. S., Stal, L. J., and Xavier, J. B. (2006). Analysis of a marine phototrophic biofilm by confocal laser scanning microscopy using the new image quantification software PHLIP. *BMC Ecol.* 6:1. doi: 10.1186/1472-6785-6-1

Pabst, B., Pitts, B., Lauchnor, E., and Stewart, P. S. (2016). Gel-entrapped *Staphylococcus aureus* bacteria as models of biofilm infection exhibit growth in dense aggregates, oxygen limitation, antibiotic tolerance, and heterogeneous gene expression. *Antimicrob. Agents Chemother.* 60, 6294–6301. doi: 10.1128/AAC.01336-16

Parker, A. E., Christen, J. A., Lorenz, L., and Smith, H. (2020). Optimal surface estimation and thresholding of confocal microscope images of biofilms using Beer's Law. *J. Microbiol. Methods* 174:105943. doi: 10.1016/j.mimet.2020.105943

Parker, A. E., Hamilton, M. A., and Goeres, D. M. (2018a). Reproducibility of antimicrobial test methods. *Sci. Rep.* 8, 1–8. doi: 10.1038/s41598-018-30282-3

Parker, A. E., Pitts, B., Lorenz, L., and Stewart, P. S. (2018b). Polynomial accelerated solutions to a large gaussian model for imaging biofilms: in theory and finite precision. *J. Am. Stat. Assoc.* 113, 1431–1442. doi: 10.1080/01621459.2017.1409121

Parker, A. E., Walker, D. K., Goeres, D. M., Allan, N., Olson, M. E., and Omar, A. (2014). Ruggedness and reproducibility of the MBEC biofilm disinfectant efficacy test. *J. Microbiol. Methods* 102, 55–64. doi: 10.1016/j.mimet.2014.04.013

Patel, D. V., and McGhee, C. N. (2013). Quantitative analysis of *in vivo* confocal microscopy images: A review. *Surv. Ophthalmol.* 58, 466–475. doi: 10.1016/j.survophthal.2012.12.003

Pettygrove, B. A., Kratofil, R. M., Alhede, M., Jensen, P. Ø., Newton, M., Qvortrup, K., et al. (2021). Delayed neutrophil recruitment allows nascent Staphylococcus aureus biofilm formation and immune evasion. *Biomaterials* 275:120775. doi: 10.1016/j.biomaterials.2021.120775

Power, R. M., and Huisken, J. (2017). A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nat. Methods* 14, 360–373. doi: 10.1038/nmeth.4224

Qin, B., Fei, C., Bridges, A. A., Mashruwala, A. A., Stone, H. A., Wingreen, N. S., et al. (2020). Cell position fates and collective fountain flow in bacterial biofilms revealed by light-sheet microscopy. *Science* 369, 71–77. doi: 10.1126/science.abb8501

Ramsey, F., and Schafer, D. (2013). The Statistical Sleuth: A Course in Methods of Data Analysis, 3rd Edn. Boston, MI: Brooks Cole.

Rani, S. A., Pitts, B., and Stewart, P. S. (2005). Rapid diffusion of fluorescent tracers into *Staphylococcus epidermidis* biofilms visualized by time lapse microscopy. *Antimicrob. Agents Chemother.* 49, 728–732. doi: 10.1128/AAC.49.2.728-732.2005

Renslow, R., Lewandowski, Z., and Beyenal, H. (2011). Biofilm image reconstruction for assessing structural parameters. *Biotechnol. Bioeng.* 108, 1383–1394. doi: 10.1002/bit.23060

Sauer, K., Camper, A. K., Ehrlich, G. D., Costerton, J. W., and Davies, D. G. (2002). *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J. Bacteriol.* 184, 1140–1154. doi: 10.1128/jb.184.4.1140-1154.2002

Schwartz, J., Leidal, K. G., Femling, J. K., Weiss, J. P., and Nauseef, W. M. (2009). Neutrophil bleaching of GFP-expressing *Staphylococci*: probing the intraphagosomal fate of individual bacteria. *J. Immunol.* 183, 2632–2641. doi: 10.4049/jimmunol.0804110

Sommerfeld Ross, S., Tu, M. H., Falsetta, M. L., Ketterer, M. R., Kiedrowski, M. R., Horswill, A. R., et al. (2014). Quantification of confocal images of biofilms grown on irregular surfaces. *J. Microbiol. Methods* 100, 111–120. doi: 10.1016/j.mimet.2014.02.020

Stewart, P. S. (2015). Antimicrobial Tolerance in Biofilms. *Microbiol. Spectr.* 3:2014. doi: 10.1128/microbiolspec.MB-0010-2014

Stewart, P. S., Davison, W. M., and Steenbergen, J. N. (2009). daptomycin rapidly penetrates a *Staphylococcus epidermidis* biofilm. *Antimicrob. Agents Chemother.* 53, 3505–3507. doi: 10.1128/AAC.01728-08

Stewart, P. S., and Parker, A. E. (2019). Measuring antimicrobial efficacy against biofilms: a meta-analysis. *Antimicrob. Agents Chemother.* 63:19. doi: 10.1128/AAC.00020-19

Thompson, S. K. (1990). Adaptive cluster sampling. *J. Am. Stat. Assoc.* 85, 1050–1059. doi: 10.1080/01621459.1990.10474975

Tilt, N., and Hamilton, M. A. (1999). Repeatability and reproducibility of germicide tests: a literature review. *J. AOAC Int.* 82, 384–389.

Tolker-Nielsen, T., and Sternberg, C. (2014). "Methods for studying biofilm formation: flow cells and confocal laser scanning microscopy," in *Pseudomonas Methods and Protocols Methods in Molecular Biology*, eds A. Filloux and J. -L. Ramos (New York, NY: Springer). doi: 10.1007/978-1-4939-0473-0_47

Voyich, J. M., Braughton, K. R., Sturdevant, D. E., Whitney, A. R., Saïd-Salim, B., Porcella, S. F., et al. (2005). Insights into mechanisms used by *Staphylococcus aureus* to avoid destruction by human neutrophils. *J. Immunol.* 175, 3907–3919. doi: 10.4049/jimmunol.175.6.3907

Voyich, J. M., Vuong, C., DeWald, M., Nygaard, T. K., Kocianova, S., Griffith, S., et al. (2009). The SaeR/S gene regulatory system is essential for innate immune evasion by *Staphylococcus aureus*. *J. Infect. Dis.* 199, 1698–1706. doi: 10.1086/598967

Vyas, N., Sammons, R. L., Addison, O., Dehghani, H., and Walmsley, A. D. (2016). A quantitative method to measure biofilm removal efficiency from complex biomaterial surfaces using SEM and image analysis. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep32694

Yang, X., Beyenal, H., Harkin, G., and Lewandowski, Z. (2000). Quantifying biofilm structure using image analysis. *J. Microbiol. Methods* 39, 109–119. doi: 10.1016/S0167-7012(99)00097-4

Yerly, J., Hu, Y., Jones, S. M., and Martinuzzi, R. J. (2007). A two-step procedure for automatic and accurate segmentation of volumetric CLSM biofilm images. *J. Microbiol. Methods* 70, 424–433. doi: 10.1016/j.mimet.2007.05.022

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Pettygrove, Smith, Pallister, Voyich, Stewart and Parker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.