STATISTICAL ANALYSIS PLAN

Study Title:	Clinical Validation Protocol for Test of Non-Inferiority of Primary Diagnosis by WSI using Hamamatsu NanoZoomer S360 Digital Slide Scanner System Compared to Conventional Determination by Light Microscopy
Study Number:	HCT-P001
Protocol Version and Date:	Ver 4, 09 JUL 2020
Sponsor:	Hamamatsu Photonics K.K.
Plan Prepared by:	Ping-Yu Liu, PhD Liu Associates Consulting, LLC
Version and Date:	Version 2.0: Date (01 AUG 2021)
Clinicaltrials.gov	NCT#: 03991468

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Term
AAP	All Available Population
AE	Adverse Event
AP	Adjudicator pathologist
CAP	College of American Pathologists
CFR	Code of Federal Regulations
CI	Confidence Interval
CMOS	Complementary Metal-Oxide Semiconductor
CRF	Case Report Form
CRO	Contract Research Organization
DPA	Digital Pathology Association
eCRF	Electronic case report form
EDC	Electronic Data Capture
EP	Enrolling Pathologist
FDA	Food and Drug Administration
FFPE	Formalin-Fixed Paraffin Embedded
GCP	Good Clinical Practice
GT	Ground Truth
H&E	Hematoxylin and Eosin
ID	Identifier
IT	Information Technology
LED	Light Emitting Diode
LM	Light Microscopy
MMRM	Mixed Model Repeated Measures
NA	Numerical Aperture
PDF	Portable Document Format
PHI	Protected Health Information
PI	Principal Investigator
PIPS	Philips IntelliSite Pathology Solution
RP	Reader Pathologist
SAP	Statistical Analysis Plan
SC	Study Coordinator
ST	Scan Technician
US	United States
USB	Universal Serial Bus
VEP	Verifying Enrolling Pathologist

Abbreviation	Term
WSI	Whole Slide Imaging

1. BACKGROUND AND RATIONALE

Whole Slide Imaging (WSI) has emerged as an alternative way to view pathology slides for primary diagnosis instead of the conventional method of viewing slides of tissue under a traditional light microscope. There is growing adoption of the use of WSI for primary diagnosis in many countries. WSI is particularly advantageous for remote consultation over vast geographical regions where pathologists are not available, and for easier archiving of images of slide material that may have a longer shelf life in digital form. In the United States, WSI is increasingly used for teaching, archiving, consultation, and research. Furthermore, the College of American Pathologists (CAP) has published recommendations to pathologists who wish to validate WSI in their clinical practice. [1]

However, quite recently, in a *de novo* authorization letter [2] and device summary [3], the Food and Drug Administration (FDA) has announced the authorization of the Philips IntelliSite Pathology Solution (PIPS) for use of WSI for primary diagnosis, specifically permitting WSI for *in-vitro* diagnostic use as an aid to the pathologist to review and interpret digital images of surgical pathology slides prepared from formalin-fixed paraffin embedded (FFPE) tissue, but not intended for use with frozen section, cytology, or non-FFPE hematopathology specimens.

Thus, the authorization of the Philips device for primary diagnosis, PIPS system, serves as the predicate device for any future 510(k) submissions for WSI systems seeking clearance for intended use of WSI, and, further, the FDA has indicated that a clinical study to prove non-inferiority of the WSI device is a "special control" required to acquire that clearance.

Hamamatsu is developing a digital slide scanner system, the NanoZoomer S360MD Digital Slide Scanner System (NanoZoomer), for the same intended use as the PIPS system. Thus, Hamamatsu will make the submission in the form of a 510(k) premarket notification with the PIPS device as the predicate device and will accordingly adhere to the special controls that were established. Hamamatsu will follow similar study designs to test the NanoZoomer system and intends to use the data from this non-significant risk study for the 510(k) submission to clear its NanoZoomer system for the same intended use. Additional study design input became available from the FDA pre-submission review of the study protocol [4, 5, 6, 7, 8, 9].

2. STUDY OBJECTIVES

2.1 Primary Objective

The primary statistical objective of this study is to demonstrate that the accuracy of the Hamamatsu NanoZoomer system is non-inferior to the diagnostic reference standard "Glass" (conventional light microscopy) in routine surgical pathology cases.

2.2 Secondary Objectives

The secondary objectives of this study are to:

1. Evaluate the accuracy of the Hamamatsu WSI test method compared to that of the reference method by site.
2. Evaluate the accuracy of the Hamamatsu WSI test method compared to that of the reference method by reader
3. Evaluate the accuracy of the Hamamatsu WSI test method compared to that of the reference method by organ.
4. Evaluate the accuracy of the Hamamatsu WSI test method compared to that of the reference method by case subtype/procedure.

3. COMPARISONS OF INTEREST AND ENDPOINT

3.1 Primary Hypothesis

Assuming a non-inferiority margin of 4%, consistent with the literature [10, 11, 12] and with the target to which the PIPS' clearance was held, then the hypothesis to be tested can be written as:

$$H_0: \pi_{WSI} - \pi_{Glass} > 0.04 \; Versus \; H_1: \pi_{WSI} - \pi_{Glass} \leq 0.04$$

where π_{WSI} and π_{Glass} are the major discordance rates for WSI and Glass respectively compared to GT.

3.2 Endpoints

3.2.1 Primary

The study's primary outcome of interest was the difference (WSI – Glass) in the rates of major discordance between the two modalities.

3.2.2 Secondary

The secondary endpoints are:

- The difference in major discordance rates between the two modalities by site
- The difference in major discordance rates between the two modalities by reader
- The difference in major discordance rates between the two modalities by organ
- The difference in major discordance rates between the two modalities by case sub-type/procedure

4. STUDY DESIGN

The study is a multi-site, randomized-read order, retrospective, paired-design evaluation of the Hamamatsu NanoZoomer S360MD Slide Scanner system consisting of a review of archived, de-identified and previously signed-out slides representing main organ systems within surgical pathology. Cases will include retrospective hematoxylin and eosin (H&E) stained formalin fixed tissue, and special stains and/or immunohistochemical stains (most commonly used is the brown chromophore) from the pathology practice, but will not include frozen sections, or cytological and hematological cases.

A total of 2000 cases consisting of multiple organ and tissue types will be enrolled, see **Table 1** for organ types, case subtype/procedure combinations, and number of cases for each organ type and case subtype/procedure. Cases will be divided over four (4) sites. At each site, each of four pathologists will read all the cases assigned to their site using both WSI and Glass modalities in a randomized order and with a washout period of at least four weeks between readings, resulting in a total of 8000 planned WSI reads and 8000 planned Glass reads. After any completed reader diagnosis CRF at a site has been collected and cleaned, two adjudicators will review the reader's diagnosis against the original diagnosis to determine whether the diagnosis was concordant, minor discordant, or major discordant compared to GT. A third adjudicator will be used if disagreement pertaining to the major discordance status occurs between the first two adjudicators.

Table 1: FDA List of Cases to be tested for Primary Diagnosis Study

(TOTAL 2000 CASES FOR THIS EXAMPLE)
CNB = Core Needle Biopsy; TUR = Transurethral Resection; LEEP = Loop Electrosurgical Excision Procedure; Dysp = Dysplasia; ECC = Endocervical Curettage; R/O ≡ Rule Out;

ORGAN	# OF CASES		SUBTYPES (procedures)
BREAST	300	50	Benign/Atypical CNB
		50	Benign/Atypical Lumpectomy
		50	In-Situ Carcinoma CNB
		50	In-Situ Carcinoma Lumpectomy
		50	Invasive Carcinoma CNB
		50	Invasive Carcinoma Lumpectomy
PROSTATE	300	120	Benign Core Bx
		30	Benign Resection
		120	Adenocarcinoma Bx
		30	Adenocarcinoma Resection
LUNG/BRONCHUS/Larynx/oral cavity/Nasopharynx	100	25	Benign/Inflammatory Bx Only
		25	Dysplasia Bx Only
		30	Carcinoma Bx
		20	Carcinoma Resection
COLORECTAL	150	50	Benign/Inflammatory Bx
		50	Adenomas Including Severe Dysp Bx
		40	Adenocarcinoma Endoscopic Bx
		10	Adenocarcinoma Resection
GE Junction	100	50	R/O Barrett's/Dysplasia Bx
		50	Non-Neoplastic/Inflammatory Bx
Stomach	100	50	Inflammatory Including R/O H. Pylori Bx
		40	Polyps/ Neoplastic Bx
		10	Polyps/ Neoplastic Resection
SKIN	175	50	Non-Neoplastic/Inflammatory Bx
		50	Squamous/Basal Cell Neoplasms Bx
		75	Melanocytic Lesions Bx
LYMPH NODE	100	75	For Presence/Absence Of Metastasis
(no micrometastases smaller than 0.5 mm)		25	Non-Neoplastic
BLADDER	100	25	Benign/Inflammatory/Non-Neo Bx
		25	Dysplasia Bx
		25	Noninvasive Carcinoma (TUR Or Bx)
		15	Carcinoma TUR/Bx

(TOTAL 2000 CASES FOR THIS EXAMPLE)
CNB = Core Needle Biopsy; TUR = Transurethral Resection; LEEP = Loop Electrosurgical Excision Procedure; Dysp = Dysplasia; ECC = Endocervical Curettage; R/O ≡ Rule Out;

ORGAN	# OF CASES			SUBTYPES (procedures)
			10	Carcinoma Resection
Gyn	150		40	Endometrial Bx/Curetting
			10	Hysterectomy for endometrial or cervical cancer
			25	Cervix Bx/Curetting (Bx, ECC)
			25	Cervix Bx/Curetting (Cone/LEEP)
			20	Ovary Benign/Non-Neoplastic
			30	Ovary Neoplastic
LIVER/BD, NEO	50		40	Core Bx
			10	Wedge Bx or Resection
Endocrine	100	ALL COMERS	50	Pancreas
			30	Thyroid
			10	Parathyroid
			10	Adrenal
BRAIN/NEURO	60		10	Non-Neoplastic
			25	Neoplastic Bx
			25	Neoplastic Resection
KIDNEY, NEOPLASTIC	50		50	All Comers (Consecutive Cases)
Salivary gland	50		50	
Hernial/Peritoneal	10		10	
Gallbladder	10		10	
Appendix	10		10	
Soft Tissue Tumors	20		20	
Anus/Perianal	50		50	Bx
Miscellaneous to reach 2000	15		15	

5. RANDOMIZATION AND BLINDING

5.1 Method of Assignment and Randomization

Randomization in this study refers to the order in which a given case will be read by a particular pathologist: Glass followed by WSI (GW) or WSI followed by Glass (WG) with at least 4 weeks between the two readings. In each site, half of the available cases will be randomly assigned to sequence GW and another half will be assigned to sequence WG. All 4 readers within each site will read both WSI and Glass for each case in that site. Randomization will occur in blocks such that there are an equal number of Glass cases and WSI cases read first within each block. The grouping of slides into containers at a site is a separate randomization to aid in operational management of the slides.

5.2 Blinding and Unblinding

Glass slides and digital images of slides scanned for each case will be de-identified and coded with a unique study identification number to ensure confidentiality. De-identification will occur per institutional policies/procedures and will follow 'honest-broker' policies or comparable institutional practices. All case identification numbers will be unlinked to patient identity and will not be individually identifiable by the reading pathologists.

Reading pathologists will also be blinded to original diagnoses.

6. SAMPLE SIZE AND POWER CONSIDERATIONS

Sample size and power considerations are based upon agreement between the diagnoses determined using Glass and WSI compared with GT.

Let $l = 1, 2, \ldots, c = 4$ index the number of sites.

Let $i = 1, 2, \ldots, n_l$ index the n_l cases/slides in site l.

Let $j = 1, 2, \ldots, r = 4$ index r pathologists who read each case from site l.

Let $k = 1, 2$ index the new ($k = 1; WSI$) and reference ($k = 2$; Glass) method.

Let u_{li} = Ground Truth diagnostic classification for the i-th case in site l.

Let y_{lijk} = diagnostic classification for the i-th case by the j-th reader with the k-th method in site l

Let $g_{lijk} = \{(y_{lijk} - u_{li})\}$ and let $a_{lijk} = \begin{cases} 1 \; if \; g_{lijk} = 0 \\ 0 \; if \; g_{lijk} \neq 0 \end{cases}$

Thus, $a_{lijk} = 1$ if y_{lijk} and u_{li} are identical (no discordance [agree] or minor discordance) and $a_{lijk} = 0$ if $y_{lijk} \neq u_{li}$ (major discordance).

Let the difference between method 1 (WSI) and method 2 (Glass) for the proportion of readers with classifications identical or minor discordance compared to the GT classification for case i in site l be denoted by

$$\mathrm{d}_{li} = \frac{\sum_{j=1}^{r}(a_{lij1} - a_{lij2})}{r}$$

Assuming $E\{d_{li}\} = \Delta$ and variance $Var\{d_{li}\} = \sigma^2$(where d_{li} is equal to average over differences of error rates between WSI and Glass compared to GT diagnosis for case *i* in site *l*) the following expression is applicable:

$$n = \left(Z_\alpha + Z_\beta\right)^2(\sigma^2) \Big/ (\delta + \Delta)^2$$

$$= \frac{\left(Z_\alpha + Z_\beta\right)^2 \pi'_{WSI}(1 - \pi'_{WSI})\pi'_{Glass}(1 - \pi'_{Glass})\{1 - \lambda + (r-1)(\xi - \eta)\}}{r(\delta + \Delta)^2}$$

For which π'_{WSI} and π'_{Glass} are the probabilities that WSI and Glass methods agree (or minor discordance) with the GT diagnosis classification retrospectively and λ = Corr(a_{lij1}, a_{lij2}) is the correlation between the classifications for the two modalities by the same reader and ξ = Corr($a_{lijk}, a_{lij'k}$) is the correlation between two readers for a given modality and η = Corr($a_{lij1}, a_{lij'2}$) is the correlation between two modalities and between two readers.

Assuming power = $(1 - \beta) = 0.90$, α = 5% two-sided, a non-inferiority margin of δ, a true difference between proportions of WSI and Glass equal to Δ in favor of Glass, number of readers r = 4 and number of sites = 4, the required sample size is provided in **Table 2** for different ranges of π'_{WSI}, Δ, λ, and (ξ–η).

Table 2: Sample Size for 4 Readers per Site and 4 Sites

π'_{WSI}	λ	(ξ–η)	$Max\ \sigma^2$ (σ^2 for $\Delta = 0$)	$\Delta = \pi'_{WSI} - \pi'_{Glass}$				
				0.0	-0.5%	-1.0%	-1.5%	-2.0%
0.6	0.5	0.150	0.114000	749	976	1326	***1904***	2968
		0.125	0.105000	690	899	1221	1754	2733
		0.100	0.09600	631	822	1116	1604	2499
	0.6	0.150	0.102000	670	874	1186	1704	2655
		0.125	0.093000	611	796	1082	1553	2421
		0.100	0.084000	552	719	977	1403	2187
0.7	0.5	0.150	0.099750	665	852	1154	1653	2568
		0.125	0.091875	604	785	1063	1522	2366
		0.100	0.084000	552	718	972	1392	2163
	0.6	0.150	0.089250	587	762	1032	1479	2298
		0.125	0.081375	535	695	941	1348	2095
		0.100	0.073500	483	628	850	1218	1893
0.8	0.5	0.150	0.076000	500	646	871	1241	1920
		0.125	0.070000	460	595	802	1143	1768
		0.100	0.064000	421	544	733	1045	1616
	0.6	0.150	0.068000	447	578	779	1111	1718
		0.125	0.062000	408	527	710	1013	1566
		0.100	0.056000	368	476	642	915	1414

This table shows, a total of 2000 cases will provide more than 90% power to demonstrate non-inferiority of WSI to Glass with a non-inferiority margin of 4% as long as the proportion of agreements with Ground Truth is at least 60%, λ is at least 0.5, (ξ-η) is at most 0.15, and $\Delta \geq$ -1.5%.

7. DETERMINATION OF PROTOCOL DEVIATIONS

Site personnel will document all protocol deviations and device deficiencies with clear explanations. Study processes, including wash out, modality read order and matching of slides with correct case information will be closely managed. A deviation will be filed for any enrolled case part that has missing, damaged or broken slides. These cases will be reviewed for consistency with the protocol across all reading pathologists. Protocol deviations will be evaluated before locking the database and unblinding the study. The deviations will be summarized and reported.

8. DISCORDANCE WITH GROUND TRUTH DIAGNOSIS

Diagnoses rendered by the Reading Pathologists at each site and captured on the EDC checklists will be compared directly to the GT diagnosis. Two adjudicators will independently review the reading to determine whether the Reading Pathologist's diagnosis and the GT diagnosis are 'concordant', 'minor discordant' or 'major discordant'. If there is a disagreement involving a major discordant status between the two adjudicators, the reading will go to a third member for further adjudication. Based on the determination of the

third member, the majority choice will be selected, or a meeting will be convened to discuss the reading and determine a consensus opinion. For the purpose of this study, the definitions of major and minor discordances are as described in Table 3:

Table 3: Definition of Discordance

Severity	Definition 13 14
Minor	No Harm: Will not result in harm No change in prognosis or a change in prognosis that is unlikely to result in a change in treatment according to standard care.
	Minimal Harm [Grade 1] Further unnecessary noninvasive diagnosis test(s) performed [e.g., blood tests or non-invasive radiological examination]. Delay in diagnosis or therapy of < 6 mos. Minor morbidity due to [otherwise] unnecessary further diagnostic effort(s) or therapy predicated on the presence of [unjustified] diagnosis.
Major	Moderate Harm [Grade 2] Further unnecessary invasive diagnostic test(s) [e.g., tissue biopsy, re-excision, angiogram, radionuclide study or colonoscopy]. Delay in diagnosis or therapy of > 6 mos. Major morbidity lasting < 6mos due to [otherwise] unnecessary further diagnostic effort(s) or therapy predicated on the presence of [unjustified] diagnosis.
	Severe Harm [Grade 3] Loss of life or limb, or other body part, or long-lasting morbidity [lasting > 6mos.].

9. ANALYSIS POPULATION

9.1 All-Available Population (AAP)

The All-Available Population (AAP) includes all cases for which at least one reader provides an evaluable outcome for either WSI or Glass modalities.

Unless otherwise specified, All Available Population (AAP) will be used for all analysis.

10. STUDY CASES

10.1 Pathologist Experience

Reader information (years of experience post-residency, average number of cases per year, type of pathologist) will be tabulated by site and overall using descriptive statistics.

10.2 Disposition of Cases

The number of cases screened, eligible and enrolled will be tabulated by site and overall. The reasons for non-inclusion in analysis will be provided by site and overall.

The deferrals (e.g., deferral to specialist, deferral to glass, etc.) will be summarized by reader, site and overall for each modality. The reason for the deferral will also be provided in a listing.

The number and percentage of cases included in AAP will be summarized by site and overall.

The number and percentage of AAP cases read by each reader will be summarized for each site, if the full number of cases is not read.

10.3 Demographics and Other Baseline Characteristics

Information about the cases to be summarized descriptively are age and sex, Sex will be summarized overall and by site as frequency and percentages. Age will be similarly summarized as mean, standard deviation, median, minimum and maximum.

10.4 Ground Truth Diagnoses

The final coded GT diagnosis obtained from the charts will be tabulated by site and overall. Information to be tabulated includes the number and percentage of cases in each of the FDA Organ, and case subtype/procedure combinations as specified in **Table 1**.

For evaluation of the GT diagnosis in the WSI modality, the reasons for re-scans of the image that were required will be tabulated by site and overall.

11. STATISTICAL ANALYSIS

11.1 Primary Objective Analysis

11.1.1 Hypothesis and Endpoint

To demonstrate that the effectiveness of WSI using Hamamatsu NanoZoomer Digital system is non-inferior to the diagnostic reference standard glass diagnosis in routine surgical pathology cases with a non-inferiority margin of 4%, the null hypothesis, H_0: $\pi_{WSI} - \pi_{Glass} > 0.04$ must be rejected in favor of the alternative hypothesis H_1: $\pi_{WSI} - \pi_{Glass} \leq 0.04$, where π_{WSI} and π_{Glass} are the major discordance rates for WSI and Glass respectively compared to GT.

11.1.2 Primary Analysis of Non-inferiority of WSI Relative to Glass

Primary Analysis Model

The primary analysis of the adjudicated comparison of each modality to the GT diagnosis will be performed using a repeated measures logistic regression model. In this model, the dependent variable is major discordance status, yes versus no. Modality (Glass or WSI), will be included in the model as fixed effect; and, site, reader, and case will be included as random effects. The AAP will be used for the primary analysis; however, any reading classified as "deferred" or as missing data will be excluded. The mixed model repeated measure (MMRM) logistic regression model can be written as:

$$\text{Ln}\left[\frac{P(Y_{ilfh}=1)}{1-P(Y_{ilfh}=1)}\right] = \mu + \alpha_i + s_l + r_{f(l)} + \beta_{h(l*j)} + \varepsilon_{ilfh} \quad (1)$$

where $P(Y_{ilfh} = 1)$ is the probability that the diagnosis is discordant with GT and Y_{ilfh} is the dichotomous outcome for case h (h = 1, 2, …., 250) , by reader f (f = 1, 2, 3, 4), within site l (l = 1, 2, 3, 4) using modality i (i = 1, 2) .

μ is the overall mean.

α_i is a fixed effect due to modality i; $\sum \alpha_i = 0$

s_l is a random effect due to site l; $s_l \sim N(0, \sigma^2_s)$, and the s_l are independent.

$r_{f(l)}$ is a random effect due to reader f nested within site l; $r_{f(l)} \sim N(0, \sigma^2_r)$, and the $r_{f(l)}$ are independent

$\beta_{h(l)}$ is a random effect due to case h nested within site l; $\beta_{h(l)} \sim N(0, \sigma^2_\beta)$, and the $\beta_{h(l)}$ are independent.

ε_{ilfh} is the random error; $\varepsilon_{ilfh} \sim N(0, \sigma^2_\varepsilon)$.

The random components are independent from each other.

Using model (1), the difference in major discordance rate ($\pi_{WSI} - \pi_{Glass}$) will be estimated and corresponding two-sided 95% confidence interval (CI) will be derived. If the upper bound of the 95% CI is less than the non-inferiority margin of 4%, then WSI will be considered non-inferior to Glass.

Supportive Analyses

The point estimates and the corresponding two-sided 95% CIs of the major discordance rates for both WSI and Glass will be estimated using model **(1)**. To support the clinical performance of NanoZoomer, the modeled major discordance rate for WSI should not exceed 7%.

In addition, the observed, unmodeled major discordance rates for the two modalities, their difference and cross tabulation will also be presented.

For the observed data, details will be provided for cases for which the Glass diagnosis was concordant (including minor discordance) with the reference diagnosis and the WSI diagnosis was a major discordance for the same reader, or vice versa. The details will include case ID, reader ID, site, organ, case sub-type/procedure, major discordance status for Glass, and major discordance status for WSI.

11.1.2.1 Analysis of Non-inferiority Including Deferred Cases

The primary endpoint analysis using the MMRM logistic regression model (1) will be repeated including deferred cases (excluding missing observations). In this analysis, deferred cases will be classified as having a minor discordance with GT.

11.1.2.2 Sensitivity Analysis

Sensitivity analyses will be performed to evaluate the robustness of the primary analysis results. To perform the sensitivity analyses, the primary analysis model (1) will be analyzed with the following assumptions for deferrals and missing data observations:

a) All deferred plus missing data observations assumed as major discordance with GT.

b) All deferred plus missing data observations for WSI assumed as major discordance with GT and All deferred plus missing data observations for Glass assumed as no major discordance with GT

11.2 Secondary Objectives Analyses

The AAP will be used for all secondary analyses. All secondary analyses will be descriptive in nature with no hypothesis testing. Missing reads and deferred reads will be excluded from these analyses.

11.2.1 Accuracy by Site, Reader, Organ, and Case Subtype/Procedure

11.2.1.1 Accuracy by Site

To describe the accuracy of WSI relative to Glass by site, the MMRM logistic model (1) with the omission of the site variable will be used. For each of the four sites, the estimated major discordance rates for the two modalities, their difference ($\pi_{WSI} - \pi_{Glass}$), and their corresponding two-sided 95% CIs will be presented in

tabular format. In addition, cross tabulation of the observed, unmodeled major discordance rates for the two modalities will be presented as well.

11.2.1.2 Accuracy by Reader

To describe the accuracy of WSI relative to Glass by reader, the MMRM logistic model **(1)** with the omission of reader and site variables will be used. The analyses will be the same as those for accuracy by site.

11.2.1.3 Accuracy by Organ

The MMRM logistic regression model **(1)** will be used for each organ. The analyses will be the same as those for accuracy by site. However, as shown in **Table 1**, sample sizes are small for some organs, e.g., hernial/peritoneal, gallbladder, appendix, etc. Should convergence not be reachable for model **(1)**, cross tabulation of the observed, unmodeled major discordance rates for the two modalities will be presented. In all cases, the observed, unmodeled major discordance rates for the two modalities and the WSI – Glass difference will be presented by organ in a tabular format for all organs.

11.2.1.4 Accuracy by Case Sub-type/procedure

As shown in **Table 1**, sample sizes are generally small for case sub-type/procedure combinations and non-convergence is likely for model **(1)**. Therefore, cross tabulation of the observed, unmodeled major discordance rates for the two modalities will be presented for each case sub-type/procedure combination.

11.3 Analysis Software

All analyses will be performed using SAS Software version 9.4.

12. REVISION HISTORY

Version	*Date*	*Description*
1.0	18 JUN 2019	Original release
2.0	01 AUG 2021	Updated to reflect revised analyses based on communication with FDA through Q-Sub process

13. REFERENCES

1. Pantanowitz L, Sinard J, Henricks W, et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Archives of Pathology & Laboratory Medicine*. Dec 2013;137(12):1710–1722.
2. FDA. IntelliSite Pathology Solution (PIPS, Philips Medical Systems). *FDA*. April 12, 2017. **https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm553358.htm**. Accessed July 28, 2018.
3. FDA. FDA Device Summary. *FDA*. October 13, 2017. **https://www.accessdata.fda.gov/cdrh_docs/pdf16/DEN160056.pdf**. Accessed July 28, 2018.
4. Holody M. Presubmission for NanoZoomer S360 Digital Slide Scanner System

5. Cui C. Re: Written Feedback for Hamamatsu NanoZoomer S360 Digital Slide Scanner System
6. Cui C. Record of Meeting for NanoZoomer S360 Digital Slide System
7. Holody M. Meeting Minutes Disagreement-Q171418/A001 – Pre-Submission Meeting Minutes for NanoZoomer S360 Digital Slide System
8. Cui C. Re: Q171418/A002 Meeting Minutes Disagreement Received
9. Cui C. Re: Q171418/A002 Review Complete
10. Bauer TW, Schoenfield l, Slaw R, Yerian L, Sun Z, Henricks W. Validation of Whole Slide Imaging for Primary Diagnosis in Surgical Pathology. *Archives of Pathology & Laboratory Medicine*. April 2013;137(4):518-524.
11. Snead DR, Tsang YW, Meskiri A, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology*. Jun 2016;68(7):1063-72.
12. Goacher E, Randell R, Williams B, Treanor D. The Diagnostic Concordance of Whole Slide Imaging and Light Microscopy: A Systematic Review. *Archives of Pathology & Laboratory Medicine*. Jan 2017;141(1):151-161.