

**Aus der Klinik und Poliklinik für Dermatologie, Venerologie und  
Allergologie  
der Universität Würzburg  
Direktorin: Prof. Dr. med. E.-B. Bröcker**

**Vergleichende Untersuchungen zur automatisierten Auswertung von  
Elispot-Platten mit verschiedenen Bildanalyse-Systemen**

**Inaugural – Dissertation  
zur Erlangung der Doktorwürde der  
Medizinischen Fakultät  
der Bayerischen Julius-Maximilians-Universität zu Würzburg  
vorgelegt von  
Alexandra Vogtmann  
aus Aschaffenburg**

**Würzburg , August 2005**

Referentin: Prof. Dr. med. E.-B. Bröcker

Koreferent: Prof. Dr. med. H. J. Girschick

Dekan: Prof. Dr. med. G. Ertl

Tag der mündlichen Prüfung: 26.04.06

Die Promovendin ist Ärztin

Für meine Eltern in Dankbarkeit.

# INHALTSVERZEICHNIS

|   |           |
|---|-----------|
| <b>1.EINLEITUNG</b>   | <b>1</b>  |
| <b>1.1.Grundbegriffe der Immunologie</b>                              | <b>1</b>  |
| 1.1.1. Die T-Zell-vermittelte Immunität                               | 2         |
| 1.1.2. Effektorfunktionen der T-Zellen                                | 3         |
| <b>1.2. Nachweismethoden Antigen-spezifischer T-Zellen</b>            | <b>6</b>  |
| 1.2.1. Lyse von Zielzellen  | 6         |
| 1.2.2. Produktion und Sekretion von Zytokinen                         | 7         |
| 1.2.3. MHC-Peptid-Tetramere   | 8         |
| <b>1.3. Der Elispot-Assay</b>   | <b>8</b>  |
| <b>1.4. Auswertung von Elispot-Proben</b>                             | <b>10</b> |
| <b>1.5. Ziele der Arbeit</b>  | <b>12</b> |
| <b>2. MATERIAL UND METHODIK</b>                                       | <b>14</b> |
| <b>2.1. Aufbau des Elispot-Assays</b>                                 | <b>14</b> |
| 2.1.1. Elisa-Prinzip  | 14        |
| 2.1.2. Elispot-Prinzip  | 15        |
| 2.1.3. Präparation der verwendeten Elispot-Proben                     | 19        |
| <b>2.2. Automatisierte Bildanalysesysteme</b>                         | <b>21</b> |
| 2.2.1. KS Elispot   | 22        |
| 2.2.2. KS Elispot compact   | 24        |
| 2.2.3. KS Elispot Software  | 27        |
| <b>2.3. Auswertungsprogramm der Elispotproben</b>                     | <b>30</b> |
| 2.3.1. Einstellung der Systemparameter zur Spoterkennung              | 30        |
| 2.3.2. Auswertung der Elispotproben                                   | 30        |
| <b>2.4. Bewertung und Darstellung von Messergebnissen</b>             | <b>34</b> |
| 2.4.1. Qualitätsparameter zur Bewertung von Messergebnissen           | 34        |
| 2.4.2. Darstellung von Ergebnissen in Methoden-vergleichenden Studien | 38        |
| 2.4.3. Darstellung der Ergebnisse mittels Bland-Altman-Plots          | 42        |
| <b>2.5. Statistische Auswertung</b>                                   | <b>43</b> |
| 2.5.1. Auswertungszeiten der Bildanalysesysteme                       | 43        |
| 2.5.2. Variabilität innerhalb der Messungen                           | 44        |
| 2.5.3. Zuverlässigkeit der Spoterkennung (Reliability)                | 44        |
| 2.5.4. Verteilung der Spotgrößen auf Maus-und Humanplatten            | 45        |
| 2.5.5. Bland-Altman-Plots   | 45        |

|   |            |
|---|------------|
| <b>3. ERGEBNISSE</b>  | <b>48</b>  |
| <b>3.1. Auswertungszeiten der Bildanalyseysteme</b>   | <b>48</b>  |
| <b>3.2. Variabilität innerhalb der Messungen</b>  | <b>50</b>  |
| <b>3.3. Zuverlässigkeit der Spoterkennung (Reliability)</b>                                 | <b>51</b>  |
| <b>3.4. Spotgrößenverteilung</b>  | <b>59</b>  |
| <b>3.5. Ergebnisse der Auswertung von Mausspots</b>   | <b>59</b>  |
| 3.5.1. Mittelwerte der Differenzen für Mausspots  | 59         |
| 3.5.2. Bland-Altman-Plots für Mausspots in absoluten Zahlen                                 | 62         |
| 3.5.3. Bland-Altman-Plots für Mausspots in relativen Zahlen                                 | 66         |
| <b>3.6. Ergebnisse der Auswertung von Mausspots in Abhängigkeit von der Spotgröße</b>       | <b>70</b>  |
| 3.6.1. Mittelwerte der Differenzen für Mausspots mit Spotdurchmesser [20-100µm]             | 70         |
| 3.6.2. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [20-100µm] in absoluten Zahlen  | 73         |
| 3.6.3. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [20-100µm] in relativen Zahlen  | 76         |
| 3.6.4. Mittelwerte der Differenzen für Mausspots mit Spotdurchmesser [100-300µm]            | 80         |
| 3.6.5. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [100-300µm] in absoluten Zahlen | 82         |
| 3.6.6. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [100-300µm] in relativen Zahlen | 84         |
| 3.6.7. Vergleich der 95%-Konfidenzgrenzen   | 88         |
| <b>3.7. Ergebnisse der Auswertung von Humanspots</b>  | <b>91</b>  |
| 3.7.1. Mittelwerte der Differenzen für Humanspots   | 91         |
| 3.7.2. Bland-Altman-Plots für Humanspots in absoluten Zahlen                                | 94         |
| 3.7.3. Bland-Altman-Plots für Humanspots in relativen Zahlen                                | 97         |
| <b>3.8. Ergebnisse der Auswertung von Humanspots in Abhängigkeit von der Spotgröße</b>      | <b>101</b> |
| 3.8.1. Bland-Altman-Plots in absoluten Zahlen   | 101        |
| 3.8.2. Bland-Altman-Plots in relativen Zahlen   | 106        |
| <b>4. DISKUSSION</b>  | <b>107</b> |
| <b>5. ZUSAMMENFASSUNG</b>   | <b>130</b> |
| <b>6. LITERATUR</b>   | <b>133</b> |

# 1. Einleitung

## 1.1. Grundbegriffe der Immunologie

Als Immunität des Organismus wird seine Fähigkeit bezeichnet, sich gegen Infektionen mit Viren und Bakterien, pathogenen Pilzen und Parasiten zu schützen und diese gezielt zu bekämpfen. Körperfremde Proteine (z.B. nach Insektenstich) und körperfremde Zellen (z.B. nach Transplantation), sowie veränderte körpereigene Zellen werden erkannt und eliminiert. Unter den Reaktionen des Immunsystems unterscheidet man die angeborene oder unspezifische Immunität von der erworbenen oder adaptiven Immunität.

Die angeborene Immunreaktion umfasst alle Prozesse, bei denen als körperfremd erkannte Zellen mittels Zytolyse und Phagozytose zerstört werden. Zu den an diesen Vorgängen hauptsächlich beteiligten Zellen zählen Makrophagen, Granulozyten, natürliche Killerzellen und das Komplementsystem. Bei Makrophagen handelt es sich um phagozytierende Zellen, die in der Lage sind, Mikroorganismen aufzunehmen und zu vernichten ohne vorher bereits Kontakt mit ihnen gehabt zu haben. Granulozyten können anhand des charakteristischen Färbeverhaltens ihrer zytoplasmatischen Granula in neutrophile, eosinophile und basophile Granulozyten eingeteilt werden. Die größte und wichtigste Gruppe stellen die neutrophilen Granulozyten dar. Sie sind zur Phagozytose, Abtötung von Keimen und Freisetzung von Entzündungsmediatoren befähigt. Den eosinophilen Granulozyten kommt eine große Bedeutung bei der Bekämpfung von parasitären Infektionen zu. Basophile Granulozyten ähneln funktionell und morphologisch den Mastzellen.

Die erworbene Immunität ist das Resultat eines Lernprozesses, während dessen das Immunsystem spezifische Antigene erkennt, analysiert und mit Bildung spezieller Proteine rasch auf diese reagieren kann. Sie ist für die lebenslange Immunität des Organismus nach Infektionen oder Impfungen verantwortlich. Die Träger der adaptiven Immunreaktion sind die B- und T-Lymphozyten.

Im Rahmen der humoralen Immunantwort erkennen die B-Zellen Antigene über den Antigenrezeptor an ihrer Oberfläche. Dieser entspricht einer membrangebundenen Form eines Antikörpers. Die zwei schweren Ketten dieses Transmembranimmunglobulins sind mit den zwei Proteinketten  $Ig\alpha$  und  $Ig\beta$  zu

einem Komplex assoziiert. Die Bindung von Antigen an den B-Zell-Rezeptor-Komplex führt zur Aktivierung der B-Zelle und ihrer klonalen Expansion unter Differenzierung zu Plasmazellen. Diese bilden Antikörper mit der Spezifität des ursprünglichen Rezeptors und sezernieren sie ins Plasma. (Janeway et al., 2002).

### **1.1.1. Die T-Zell-vermittelte Immunität**

Die T-Lymphozyten gelten als die Vermittler der zellulären Immunantwort. Für die Antigenerkennung durch T-Zellen ist charakteristisch, dass sie die Anwesenheit intrazellulärer Pathogene nur dann erkennen, wenn ein Peptidfragment der Erregerantigene als Komplex aus Peptid und MHC-Molekül (major histocompatibility complex, Haupthistokompatibilitätskomplex) auf der Oberfläche der infizierten Zelle präsentiert wird. Die MHC-Moleküle werden in zwei Klassen unterteilt. Die MHC-Klasse-I-Moleküle transportieren Peptide von sich im Zytoplasma vermehrenden Pathogenen (z.B. Viren und bestimmte Bakterien) an die Zelloberfläche. Sie werden auf allen kernhaltigen Zellen exprimiert. Die MHC-Klasse-II-Moleküle binden Peptide von Pathogenen, die sich in intrazellulären Vesikeln vermehren, sowie Peptidfragmente von aufgenommenen extrazellulären Erregern, Toxinen und Tumor-assoziierten Antigenen. MHC-Klasse-II-Moleküle befinden sich auf Antigen-präsentierenden Zellen (Makrophagen, dendritische Zellen, B-Zellen).

Die menschliche MHC-Klasse ist polygen und hochpolymorph. Es existieren mehrere MHC-Klasse-I- und MHC-Klasse-II-Gene. Für jedes der Gene gibt es multiple Allele. Die Vielfalt der MHC-Glykoproteine ist für die Antigenerkennung durch T-Zellen von großer Relevanz, da eine T-Zelle ein Peptid nur erkennt, wenn es als Komplex mit einer bestimmten allelen Variante des MHC-Moleküls vorliegt (MHC-Abhängigkeit der T-Zelle).

Unter den T-Zellen unterscheidet man die CD8<sup>+</sup>-T-Lymphozyten und die CD4<sup>+</sup>-T-Zellen. CD8<sup>+</sup>-T-Lymphozyten erkennen Peptide, die an MHC-Klasse-I-Moleküle gebunden sind. Sie differenzieren nach Antigenerkennung zu zytotoxischen T-Zellen, die infizierte Zellen zerstören können.

Die CD4<sup>+</sup>-T-Zellen (TH-Zellen) üben eine immunmodulatorische und aktivierende Wirkung auf andere Immunzellen aus. Peptidfragmente der Erregerantigene

werden den CD4<sup>+</sup>T-Zellen in Verbindung mit MHC-Klasse-II-Molekülen auf der Oberfläche von Antigen-präsentierenden Zellen gezeigt. Die CD4<sup>+</sup>T-Zellen können daraufhin in zwei funktionell unterschiedliche Gruppen differenzieren, die TH1- und TH2-Zellen.

Die Induktion zu TH1-Zellen (inflammatorische T-Zellen) erfolgt v. a. in Anwesenheit von Interleukin (IL)-12, das von aktivierten Makrophagen und dendritischen Zellen produziert wird. Die TH1-Zellen stimulieren die keimtötenden Eigenschaften von Makrophagen und veranlassen B-Zellen zur Produktion von IgG-Antikörpern.

Für die Entwicklung der TH2-Zellen (T-Helferzellen) ist v. a. IL-4 bedeutsam. Die TH2-Zellen besitzen die Fähigkeit B-Zellen zur IgM-Antikörperproduktion anzuregen und fördern damit die humorale Immunreaktion. Sie aktivieren auch die IgE und IgA-Antikörperbildung.

### **1.1.2. Effektorfunktionen der T-Zellen**

Die entscheidenden Vermittler der Effektorfunktionen der CD4<sup>+</sup>Zellen sind Zytokine. Sie wirken als Botenstoffe von Immunzellen. Zu den Zytokinen zählen Interferone, Interleukine und bestimmte Wachstumsfaktoren. Die kleinen Proteine können gelöst oder als membrangebundene Effektormoleküle vorliegen. Sie binden an spezifische Rezeptoren auf einer Zielzelle und vermitteln Aktivierungs-, -Proliferations- und Differenzierungsinformationen. Zu den wichtigsten Zytokin-bildenden Zellen zählen aktivierte Makrophagen und T-Zellen.

TH2-Zellen sezernieren die Interleukine (IL) IL-4, IL-5 und IL-10, die auf B-Zellen aktivierend wirken. IL-4 begünstigt die IgE-Produktion, IL-5 hat einen aktivierenden Effekt auf eosinophile Granulozyten.

Die antiinflammatorischen Zytokine IL-4, IL-10, IL-13 und TGF- $\beta$  (transforming growth factor, transformierender Wachstumsfaktor) wirken hemmend auf die Aktivierung von Makrophagen durch IFN- $\gamma$ .

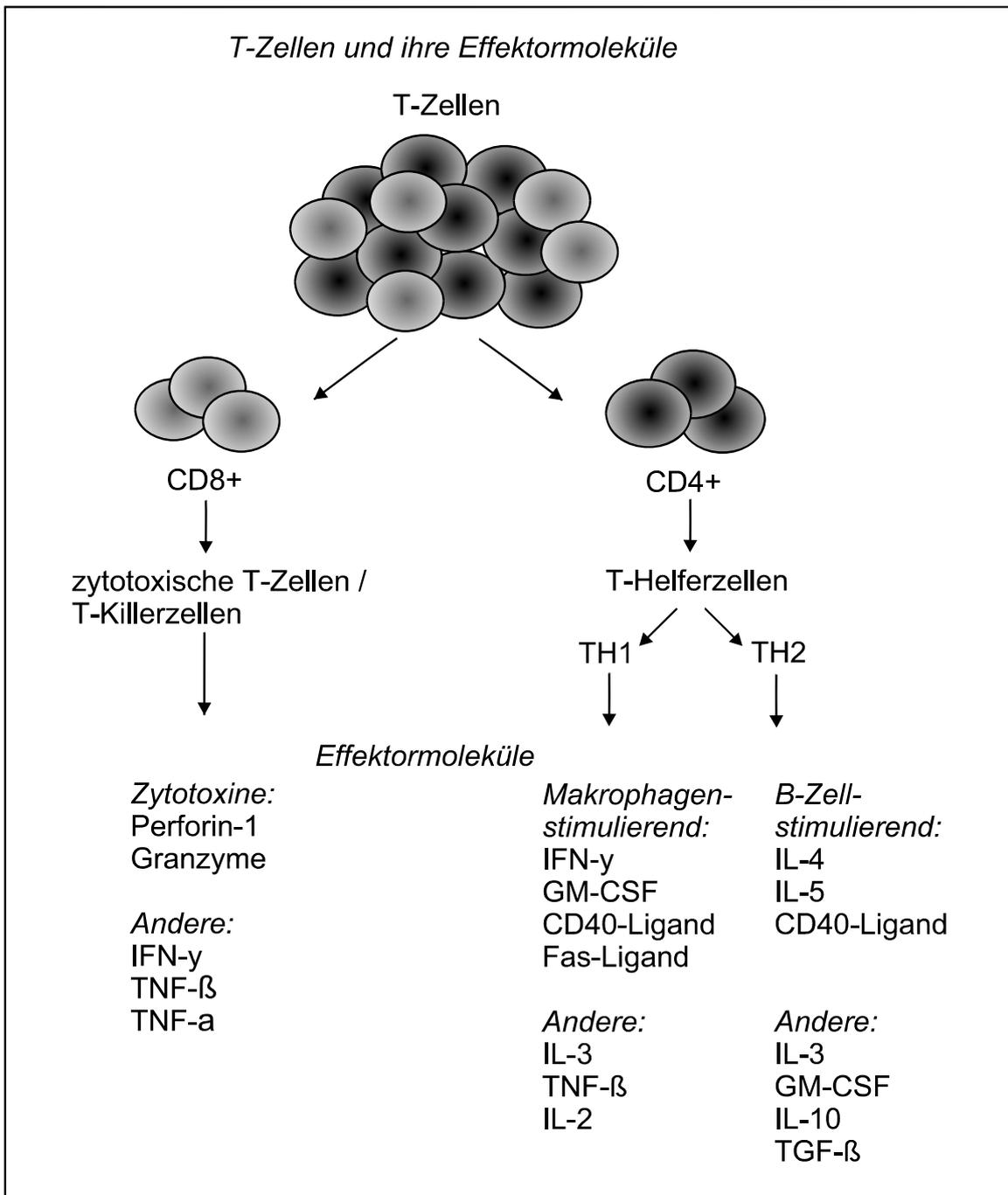
Das charakteristische Zytokin der TH1-Zellen ist IFN- $\gamma$ . Es stellt das wichtigste makrophagenaktivierende Zytokin dar und hat eine stimulierende Wirkung auf die Produktion der IgG-Antikörper, die extrazelluläre Pathogene effektiv opsonieren und so ihre Aufnahme durch phagozytierende Zellen unterstützen.

IFN- $\gamma$  fördert zusammen mit IL-2 die Differenzierung der CD8<sup>+</sup>-T-Lymphozyten zu zytotoxischen T-Zellen.

TH1-Zellen sezernieren zudem TNF- $\alpha$  (tumor necrosis factor, Tumornekrosefaktor) und TNF- $\beta$ , die ebenfalls makrophagenstimulierende Effekte aufweisen. Sie exprimieren den CD40-Liganden und/oder den Fas-Liganden, der Apoptose induziert.

Von beiden Typen der CD4<sup>+</sup>T-Zellen werden die auch entfernt wirksamen Zytokine IL-3 und GM-CSF (Granulozyten-Makrophagen-Koloniestimulierender Faktor) produziert. Sie stimulieren die Myelopoese im Knochenmark und regen die Bildung von dendritischen Zellen aus Vorläuferzellen an.

Die Haupteffektormoleküle der CD8<sup>+</sup>-T-Lymphozyten sind die Zytotoxine. Proteolytisch wirksame Granzyme, porenbildendes Perforin und der Apoptose induzierende Fas-Ligand werden in speziellen Granula gespeichert und sezerniert. Die CD8<sup>+</sup>-T-Zellen bilden außerdem die Zytokine IFN- $\gamma$ , TNF- $\alpha$ , TNF- $\beta$ . (Janeway et al., 2002), (Abbas and Lichtman, 2004).



**Abb. 1:** Die drei Arten von T-Effektorzellen (TH1-Zellen, TH2-Zellen, zytotoxische T-Lymphozyten) mit den verschiedenen von ihnen produzierten Effektormolekülen. (Modifiziert nach Janeway et al., 2002.)

## **1.2. Nachweismethoden Antigen-spezifischer T-Zellen**

Zum Nachweis Antigen-spezifischer T-Zellen existieren heute mehrere Verfahren, die auf unterschiedlichen Prinzipien beruhen (Romero et al., 1998), (Speiser et al., 2004). Antigen-spezifische T-Zellen können zum einen über die Zytotoxin-vermittelte Lyse von Zielzellen erfasst werden. Zu den auf diesem Prinzip basierenden Verfahren zählen der  $^{51}\text{Cr}$ -Release-Assay und der Limiting-Dilution-Assay. Zum anderen erfolgt der Nachweis Antigen-spezifischer T-Zellen über die von ihnen produzierten und sezernierten Zytokine mittels des Sandwich-Elisa (enzyme linked immunosorbent assay), des Elispot-Assays (enzyme linked immunospot assay) oder der intrazellulären FACS (fluorescence activated cell sorting)-Analyse. Die dritte Nachweismethode zur Detektion von Antigen-spezifischen T-Zellen stellen die MHC-Peptid-Multimere dar. Den funktionellen Nachweismethoden zugeordnet werden dabei der Sandwich-Elisa, der Elispot-Assay, der Chromium-Release-Assay sowie der Limiting-Dilution-Assay, während die MHC-Peptid-Tetramere und die FACS-Analyse als strukturelle Nachweisverfahren gelten.

### **1.2.1. Lyse von Zielzellen**

Ein Verfahren, das Antigen-spezifische  $\text{CD8}^+$ -T-Zellen mittels ihrer zytotoxischen Wirkung auf Zielzellen nachweist, ist der  $^{51}\text{Cr}$ -Release-Assay. Mit dem Isotop Chromium ( $^{51}\text{Cr}$ ) inkorporierte Zielzellen werden mit einem spezifischen Antigen und T-Zellen für mehrere Stunden inkubiert. Nach erfolgreicher Antigenerkennung sezernieren die T-Zellen Zytotoxine, die zur Lyse der Zielzellen und damit zur Freisetzung des radioaktiven Chromiums führen, welches mit speziellen Messgeräten quantifiziert werden kann. Dabei besteht eine Korrelation zwischen der freigesetzten Menge an radioaktivem Chromium und dem Grad der Zytotoxizität. Der biologische Effekt der Zelllyse wird über die, von spezifischen T-Zellen nach Antigenerkennung freigesetzten Zytotoxine wie z.B. Perforin und Granzym vermittelt.

Der Limiting-Dilution-Assay (LDA), auch als limitierender Verdünnungstest bezeichnet, ist eine weitere funktionelle Nachweismethode, bei der spezifische

T-Zellen mittels Lyse von Zielzellen quantifiziert werden können. In die Vertiefungen einer 96-Well-Platte werden T-Zellen, das spezifische Antigen, sowie Stimulatorzellen eingebracht. Diese Methode setzt eine anschließende 10 bis 14-tägige Expansion der T-Zellen *in vitro* voraus, was sie somit langwierig und anfällig für Störfaktoren macht. Nach mehreren Tagen wird die spezifische Antigenreaktion in Form der Destruktion von Zielzellen abgelesen. Der Logarithmus des Anteils der Vertiefungen ohne Reaktion wird gegen die Zahl der ursprünglich in die Vertiefung pipettierten Zellen aufgetragen. Mittels Anwendung der Poisson-Verteilung kann daraus der Anteil der antigenspezifischen Zellen in der Population bestimmt werden.

### **1.2.2. Produktion und Sekretion von Zytokinen**

Ein Verfahren, das die von Antigen-spezifischen T-Zellen freigesetzten Zytokine in Zellkulturflüssigkeiten nachweist, ist der Sandwich-Elisa (enzyme linked immunosorbent assay). Die an einen Träger fixierten Antizytokin-Antikörper binden das in der zu untersuchenden Flüssigkeit vorhandene Zytokin und konzentrieren es an der Trägersoberfläche. Ein Enzym-gekoppelter zweiter Antikörper bindet an diese Komplexe. Nach Zugabe des Enzymsubstrats wird die in der Enzym-Substrat-Reaktion entstandene Produktmenge quantifiziert.

Der Elispot-Assay (enzyme linked immunospot assay) stellt ein modifiziertes Elisa-System dar. Hierbei werden spezifische T-Zellen, die nach Antigenerkennung Zytokine produzieren, auf Einzelzellebene erfasst. Im Unterschied zum Elisa, der die Zytokinfreisetzung als Konzentration quantitativ in einer Zellflüssigkeit misst, weist der Elispot-Assay die einzelne, spezifische T-Zelle über die von ihr sezernierten Zytokine nach.

Die intrazelluläre Zytokin FACS (fluorescence activated cell sorting)-Analyse identifiziert spezifische T-Zellen mittels Verwendung der Durchflusszytometrie und der neuen Methoden zur intrazellulären Anfärbung von Zytokinen. Die Zellen werden mit einem fluoreszenzmarkierten, monoklonalen Anti-Zytokin-Antikörper gefärbt, strömen dann einzeln an einem Fotodetektor vorbei und können anhand des jeweiligen Zytokinexpressionsmusters charakterisiert werden.

### **1.2.3. MHC-Peptid-Tetramere**

Eine neuere Methode mit der Antigen-spezifische T-Zellen detektiert werden können, sind die MHC-Peptid-Tetramere-und Multimere. MHC-Moleküle werden mit spezifischen Antigenen bestückt, mit Streptavidin tetramerisiert und mit fluoreszierenden Substanzen markiert. Das Streptavidinmolekül bindet über seine vier (→tetra) Biotinketten die mit Peptid bestückten MHC-Komplexe. Nach Bindung dieser Tetramere an den T-Zellrezeptor können die spezifischen T-Zellen mittels Durchflusszytometrie quantifiziert werden. Die Methode weist die Antigen-spezifischen T-Zellen direkt nach. Darüber hinaus ist das Verfahren geeignet, Antigen-spezifische T-Zellen auch in situ zu detektieren (Berger et al., 2004). Es erlaubt jedoch nicht, Aussagen zu deren Funktionalität zu treffen.

### **1.3. Der Elispot-Assay**

Die Elispot (enzyme linked immunospot assay)-Technik wurde 1983 ursprünglich zum Nachweis von Antikörper-produzierenden B-Zellen entwickelt (Czerkinsky et al., 1983). Als „reverse Elispot assay“ wurde die Methode 1988 zur Erfassung von einzelnen Lymphokin-sezernierenden Zellen aus dem peripheren Blut etabliert (Czerkinsky et al., 1988). Herr et al. ermöglichte dann erstmals den Einsatz der Elispot-Technik zum Nachweis von tumor-und antigenspezifischen T-Zellen. Heute findet der Elispot-Assay auf vielen medizinischen Gebieten breite Anwendung und wird in zahlreichen Publikationen als verwendetes Nachweisverfahren antigenspezifischer T-Zellen angegeben.

Der Elispot-Assay wird als „Immunmonitoring“ in der Erforschung und Anwendung immunologischer Therapieansätze bei der Behandlung von Tumor- und Autoimmunerkrankungen angewandt. (Coulie, 1997), (Wang et al., 1999), (Gajewski, 2000), (Eming et al., 2000), (Farina et al., 2001), (Griffioen et al., 2001), (Hellings et al., 2001), (Andersen et al., 2001), (Schaed et al., 2002), (Kiessling et al., 2002), (Otto et al., 2005).

Er hat einen bedeutenden Einsatz in der Entwicklung von Impfstoffen und Behandlungsstrategien gegen chronische Virusinfektionen, darunter AIDS, (Lee et al., 1989), (Corne et al., 1999), Hepatitis, (Anthony et al., 2001), HPV-, (Davidson et al., 2003), (van der Burg et al., 2001), EBV-, (Yang et al., 2000),

CMV-, (Khan et al., 2002), (Godard et al., 2004), Herpes-simplex-, (Schmid et al., 1997a), Varizella-zoster-Virus-Infektionen (Smith et al., 2003) sowie gegen bakterielle Infektionen, z.B. mit *Helicobacter pylori* (Bontems et al., 2003), *M. tuberculosis* (Lalvani et al., 2001), (Ewer et al., 2003), (Hill et al., 2005) und *Treponema pallidum* (Tabidze et al., 1999).

Darüber hinaus wird das Verfahren zur Diagnostik allergischer Erkrankungen und zur Verlaufskontrolle von Therapien, wie z.B. spezifischen Hyposensibilisierungen, angewandt. (Gabrielsson et al., 2001), (Lindemann et al., 2003). In der Routineanwendung ermöglicht es eine standardisierte Beurteilung der Immunlage und der Immunreaktionen von Patienten im Verlauf von chronischen Erkrankungen. Auf dem Gebiet der Transplantationsmedizin wird die Methode zur Überwachung von Abstoßungsreaktionen eingesetzt (Hricik et al., 2003), (Tambur et al., 2005).

Die Einführung der Elispot-Methode als Nachweisverfahren für antigenspezifische T-Zellen im Jahr 1988 fand bei vielen Wissenschaftlern große Beachtung. Sie arbeiteten seitdem intensiv daran, die einzelnen Komponenten des Assays zu verbessern und bestmöglich aufeinander abzustimmen. So wurden die Mechanismen der Antigenpräsentation untersucht und gezielte Strategien zu deren Optimierung entwickelt (Britten et al., 2002), (Thumann et al., 2003). Geeignete Antizytokin-Antikörper wurden erforscht und deren notwendige Konzentrationen ermittelt. Auch andere verwendete Materialien, z.B. Elispotplatten, Enzyme, Substrate, wurden an ihre Anwendung angepasst und weiterentwickelt.

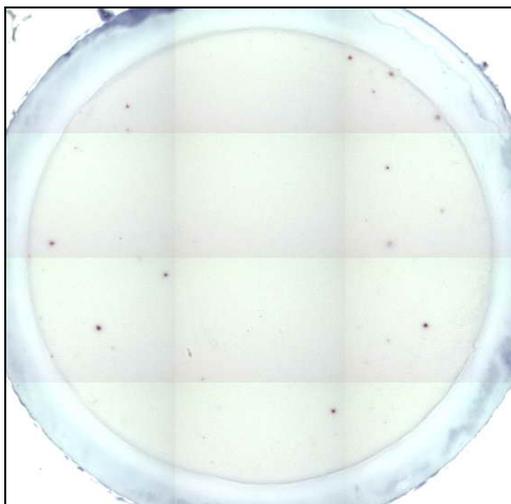
Durch diese zunehmende Standardisierung und kontinuierliche Validitätsprüfung konnte das Elispot-Verfahren immer weiter optimiert werden, so dass es heute bevorzugt zur qualitativen und quantitativen Charakterisierung von Immunprozessen eingesetzt wird. (Clay et al., 2001), (Mwau et al., 2002). Im Vergleich mit anderen Methoden besitzt der Elispot-Assay eine hohe Sensitivität zur Detektion spezifischer Immunzellen und ist als eine zuverlässige und reproduzierbare Methode anerkannt. (Kabilan et al., 1990), (Miyahira et al., 1995), (Herr et al., 1996), (Schmittel et al., 1997), (Scheibenbogen et al., 1997a), (Scheibenbogen et al., 1997b), (Pass et al., 1998), (Romero et al., 1998), (Schmittel et al., 2000), (Lewis et al., 2000), (Carvalho et al., 2001), (Bennouna et al., 2002), (Mashishi and Gray, 2002).

Das Spektrum der analysierbaren Sekretionsprodukte von Zellen mittels Elispot wächst stetig. So können derzeit TNF $\alpha$ -, TNF $\beta$ -, IL 1-, IL 2-, IL 5-, IL 6-, IL-8-, IL 10-, IL-12-, IL-13-, IL-16-, MIG-, RANTES-, latent TGF- $\beta$ 1- und GM-CSF-produzierende Zellen bei verschiedenen Spezies (Mensch, Maus, Kaninchen, Ratte, Feline, Porcine, Primate) gemessen werden. (Nordstrom and Ferrua, 1992), (Merville et al., 1993), (Fujihashi et al., 1993), (Schmid et al., 1997b), (Medhat et al., 1998), (Kumar et al., 2001).

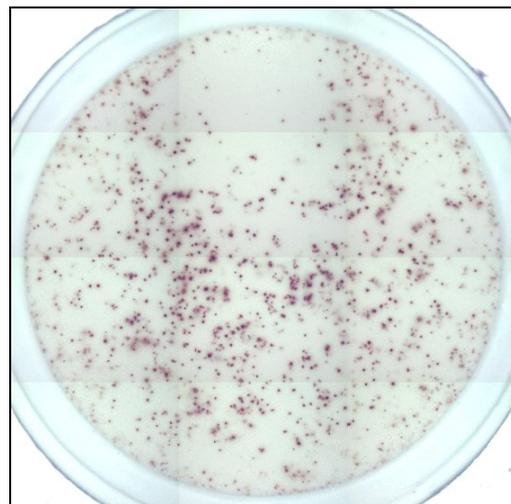
Während in den letzten Jahren nachhaltig und erfolgreich an der technischen Weiterentwicklung der Elispot-Methode gearbeitet wurde, erlangten die Standardisierung und Validitätskontrolle der Auswertung von Elispot-Proben bislang weit weniger Interesse.

#### **1.4. Auswertung von Elispot-Proben**

Der Elispot-Assay stellt positive, spezifische T-Zell-Reaktionen in Form kleiner Farbpunkte, so genannter Spots, mit einem Durchmesser von ca. 20 bis 300 $\mu$ m, dar, die sich auf dem Boden einer Vertiefung (Well) mit einem Durchmesser von 6mm befinden.



**Abb. 2a**



**Abb. 2b**

**Abb. 2a, b:** Beispiele für eine Vertiefung (Well) der Elispotplatte mit wenigen Spots (Abb. 2a) und mit sehr vielen Spots (Abb. 2b).

Die Spots auf der Elispot-Platte wurden zunächst nur mittels Stereomikroskop (10-fache Vergrößerung) oder per Auge ausgezählt. Im idealen Fall liegen die Spots einzeln, klar voneinander getrennt und gleichmäßig verteilt auf dem Boden der Vertiefung. Befinden sich im Auswertungsfeld jedoch sehr viele Spots (>500) oder überlappen die Farbpunkte, ist es dem Untersucher kaum möglich eine korrekte Anzahl anzugeben, sowie mögliche Artefakte von wahren Spots zu unterscheiden. Unter diesen Bedingungen bleibt das Ergebnis oftmals nur eine Schätzung der Spotanzahl.

Das Auszählen der Elispot-Platten per Auge oder mittels Stereomikroskop variiert zudem in Abhängigkeit vom Untersucher. Jeder Auswerter bestimmt die Kriterien, anhand derer er einen Farbpunkt als Spot klassifiziert und in die Zählung aufnimmt, während er andere verwirft und nicht ins Endergebnis einbezieht, individuell und für jede Vertiefung neu. Die Ergebniserfassung mittels Stereomikroskop oder nur mit Auge ist demzufolge subjektiv und damit nicht in ausreichend hohem Maß objektivierbar und reproduzierbar. Darüber hinaus ist sie sehr zeit- und personalaufwendig.

1997 wurde das erste automatisierte bildanalytische Auswertungsverfahren für Elispot-Platten von der Firma Zeiss entwickelt. (Cui and Chang, 1997), (Herr et al., 1997).

Mittlerweile wurden für die objektive Auswertung von Elispot-Proben verschiedene automatisierte Elispot-Reader entwickelt. Die Systeme besitzen als wesentliche Komponenten ein Mikroskop, eine Kamera, die die Spots aufnimmt und einen PC mit spezieller Software, die die Auswertung der digitalisierten Bilder ermöglicht. Bei manchen Systemen gelangt das Well-Bild über ein Mikroskop in die elektronische Datenverarbeitung, während es bei anderen Systemen direkt von der Videokamera aufgenommen wird. Der Hauptunterschied zwischen den Systemen liegt in der Mikroskopoptik und der Auflösung, mit der die Bilder aufgenommen werden. Gerade bei kleinen Spots hat die Höhe der Auflösung entscheidende Auswirkung darauf, ob ein Spot erkannt oder nicht erkannt wird. Weiterhin unterscheiden sich die einzelnen Geräte in ihrer Bedieneroberfläche und der Aufwendigkeit ihrer Handhabung, in der zur Auswertung benötigten Zeit, sowie in den Anschaffungskosten.

Angesichts der Schwächen einer Auszählung der Elispot-Assays mit Auge oder Stereomikroskop, verwenden immer mehr Arbeitsgruppen zur Auswertung ihrer

Elispot-Platten automatisierte Elispot-Reader-Systeme. Bislang wurden jedoch kaum Anstrengungen unternommen, die Qualität dieser Auswertungsmethode zu überprüfen.

Eine Arbeitsgruppe am Memorial-Sloan Kettering Cancer Center (MSKCC) in New York, USA, (Janetzki et al., 2004), verglich als erste die verschiedenen Evaluationsmethoden von Elispot-Platten, insbesondere beschäftigte sie sich mit dem Einfluss der Auswertungsmethode und dem subjektiven Einfluss des Auswerters auf die Variabilität der Ergebnisse. Im Rahmen dieser Vergleichsstudie werteten drei Wissenschaftler Elispotproben unter drei unterschiedlichen Bedingungen aus: mit dem Stereomikroskop, mit einem automatisierten bildanalytischen Auswertungssystem, an dem die Untersucher die Parameter zur Spoterkennung selbst einstellten und am automatisierten bildanalytischen Auswertungssystem, an dem die Spoterkennungsparameter bereits eingestellt waren und von den Untersuchern nicht verändert wurden. Es konnte gezeigt werden, dass die manuelle Auswertung von Elispotproben mit dem Stereomikroskop die größte Variabilität aufwies und die Auswertung deutlich abhängig vom Untersucher war, sowohl innerhalb dieser Auswertungsmethode als auch im Vergleich mit den zwei anderen Methoden. Die Entwicklung der automatisierten Lesesysteme bedeutete somit einen weiteren großen Fortschritt in der Standardisierung und Reproduzierbarkeit der Elispot-Methode.

### ***1.5. Ziele der Arbeit***

Die o. g. Vorarbeit (Janetzki et al., 2004) konnte belegen, dass die automatisierte Auswertung der Elispot-Proben die manuelle Auswertung am Stereomikroskop oder mit Auge deutlich an Objektivierbarkeit, Standardisierbarkeit und Reproduzierbarkeit übertrifft.

Unter der Vielzahl der angebotenen Elispot-Reader-Systeme bestehen große Unterschiede in der Bildauflösung, Bildaufnahme und Bildverarbeitung. Es ist anzunehmen, dass auch die automatisierten Lese-Systeme mit unterschiedlicher Qualität arbeiten. Bislang sind jedoch noch keine Untersuchungen durchgeführt worden, die die eventuellen Unterschiede zwischen den Auswertungssystemen belegen, quantifizieren und einer Ursache zurückführen könnten.

Um diese Fragestellungen zu beantworten, sollten in dieser Arbeit unterschiedliche automatisierte Bildanalyzesysteme zur Auswertung von Elispotproben miteinander verglichen werden.

Die Arbeit sollte einen Beitrag zur Verbesserung der bislang wenig beachteten Auswertung von Elispot-Proben und damit zur weiteren Standardisierung und Optimierung der Qualität der Elispot-Technik leisten.

Es wurden zwei automatisierte Elispot-Reader mit drei unterschiedlichen Auflösungsvermögen für den Vergleich herangezogen. An den beiden Systemen sollten mehrere Elispotproben von verschiedenen Spezies und unterschiedlichen, unabhängigen Arbeitsgruppen ausgewertet werden. Neben dem Vergleich der Anzahl der gemessenen Spots, sollten die Zeit für die Messungen, die Genauigkeit und Zuverlässigkeit der Messungen, sowie die Einflussgrößen der einzelnen Bildanalyzesysteme untersucht werden. Der Vergleich der Systeme sollte in einer Form durchgeführt werden, die die, in der täglichen Praxis der Elispotprobenauswertung herrschenden Abläufe optimal imitiert und es somit erlaubt, objektive Informationen für die automatisierte Elispot-Assay-Auswertung zu geben. Die Anwender der Elispot-Methode sollten aus den Ergebnissen dieser Arbeit die für ihre Bedürfnisse geeigneten automatisierten Reader-Systeme mit dem notwendigen Auflösungsvermögen entnehmen können.

## **2. Material und Methodik**

### **2.1. Aufbau des *Elis*pot-Assays**

Die *Elis*pot-Methode basiert auf dem Prinzip des Enzymimmunoassays „*Elis*a“ (enzyme linked immunosorbent assay). Im Rahmen eines *Elis*a werden Enzyme an Antikörper oder an Antigene gebunden, ohne dadurch die immunologischen oder enzymatischen Eigenschaften der Reaktionspartner zu verändern. Durch Zugabe eines Substrats und Bestimmung dessen enzymatischer Umsetzung können Antikörper-Antigen-Reaktionen sichtbar gemacht und quantitativ erfasst werden.

#### **2.1.1. *Elis*a-Prinzip**

Im ersten Arbeitsschritt werden auf einer festen Phase, z. B. einer Mikrotiterplatte, der Antikörper (AK) gegen ein gesuchtes Antigen (AG) (oder das Antigen gegen einen gesuchten Antikörper) aufgetragen. Nach Inkubation mit der Testlösung bindet dieser Antikörper (das Antigen) das spezifische gesuchte Antigen (den spezifischen AK). In einem Waschschrift werden nicht gebundene Reaktionspartner entfernt.

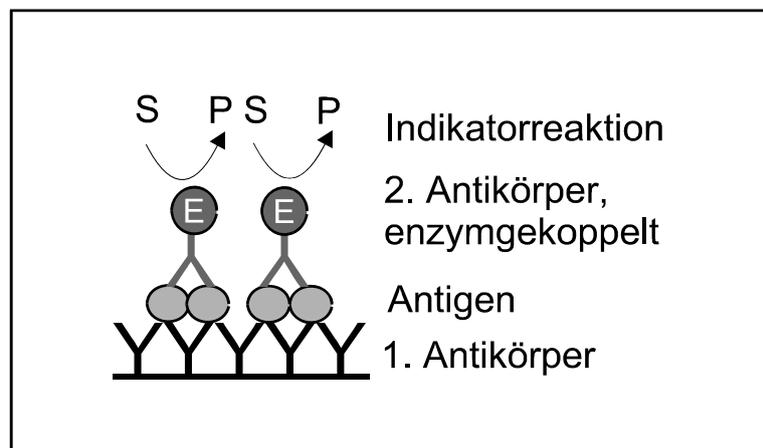
Im zweiten Arbeitsgang wird ein Enzym-gekoppelter Antikörper hinzugegeben, der gegen das gesuchte Antigen (gegen den gesuchten Antikörper) gerichtet ist und an die bereits entstandenen AK-AG-Komplexe bindet. Nicht gebundene Reaktionspartner werden in einem Waschschrift entfernt.

Im letzten Arbeitsprozess wird das Substrat des Enzyms hinzugefügt. In der Enzym-Substrat-Reaktion entsteht ein Produkt, dessen Menge gemessen wird. So kann z. B. nach Zugabe von Chromogen der gebildete Farbstoff photometrisch quantifiziert werden und dessen Menge in standardisierten Tests als proportional zur Antikörper-Antigen-Reaktion angesehen werden. Je nach Zusammenstellung der Reaktionspartner, kann jeweils nach spezifischen Antikörpern oder Antigenen gesucht werden.

Im indirekten *Elis*a wird zum Nachweis von Antikörpern das Antigen an die feste Phase gebunden, anschließend zuerst das Testserum, dann der Enzym-

gekoppelte Antikörper und schließlich das Substrat hinzugefügt. Die in der Enzym-Substrat-Reaktion gebildete Produktmenge wird quantifiziert. Diese Testanordnung wird z. B. beim Test auf Antikörper gegen das HI- Virus 1 und 2, sowie auf zahlreiche andere virale Antikörper verwendet.

Der Sandwich-Elisa, eine Doppel-Antikörper-Methode, dient dem Nachweis von gesuchten Antigenen. In dieser Anordnung bindet ein gegen das Antigen gerichteter Antikörper an die feste Phase. Nach Inkubation mit dem Testserum wird ein zweiter Enzym-markierter Anti-Antigen-Antikörper hinzugegeben. Nach Inkubation mit dem Enzymsubstrat erfolgt die Bestimmung der entstandenen Produktmenge (Abb. 3).



**Abb. 3:** Prinzip des Sandwich-Elisa. Der an die feste Phase gebundene Antikörper bindet nach Inkubation mit Testserum das spezielle Antigen. An diese Antigen-Antikörper-Komplexe bindet ein zweiter Enzym-gekoppelter Antikörper (E). Nach Zugabe des Substrats (S) wird die in der Indikatorreaktion entstandene Produktmenge (P) quantifiziert.

### 2.1.2. Elispot-Prinzip

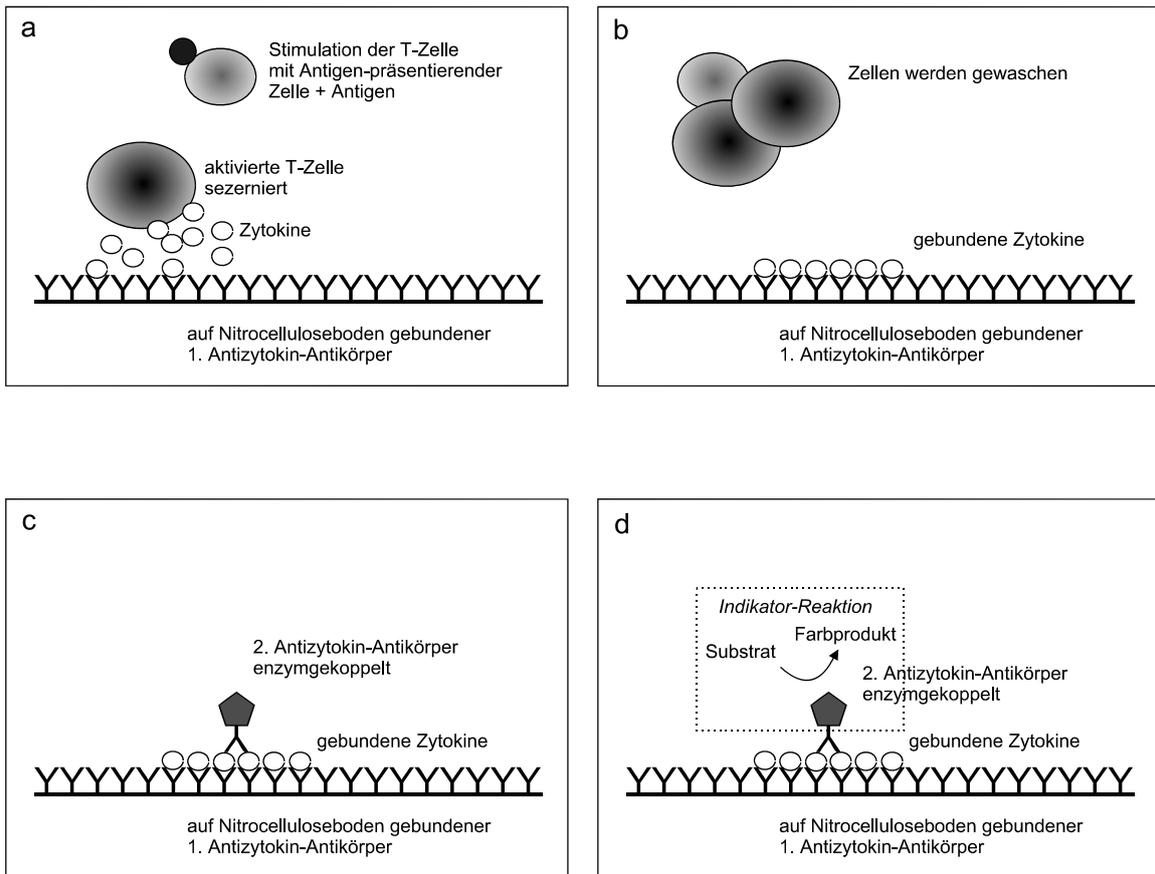
Der Elispot-Assay (enzyme linked immunospot assay) stellt ein modifiziertes Elisa-System dar. Im Unterschied zum Elisa, der quantitativ die von T-Zellen sezernierte Zytokinmenge im Serum oder in einer Zellkulturflüssigkeit nachweist,

ermöglicht die Elispot-Methode die Erfassung antigenspezifischer T-Zellen auf Einzelzellebene, die nach Antigenerkennung bestimmte Zytokine produzieren.

Die Nitrocelluloseböden einer Mikrotiterplatte mit 96 Vertiefungen (96-Well-oder 96-Loch-Mikrotiterplatte), werden mit monoklonalen Antikörpern gegen ein spezifisches, zu detektierendes Zytokin über Nacht beschichtet. Die nicht am Boden haftenden Antikörper werden in einem Waschschrift am folgenden Tag entfernt, die übrigen mit Serum geblockt.

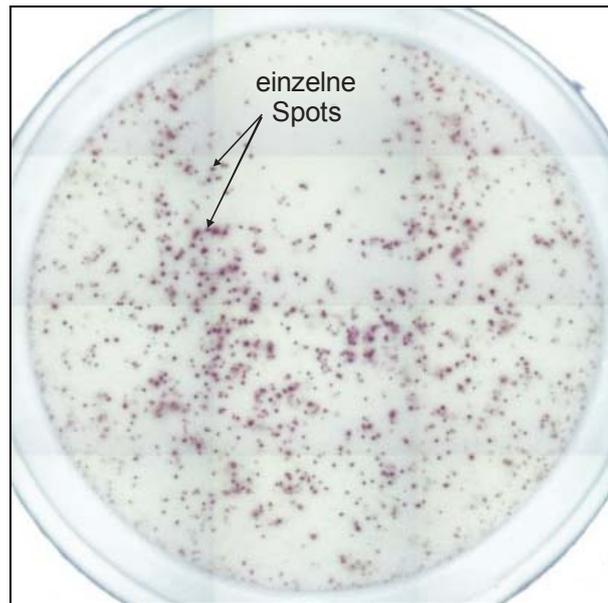
In die Vertiefungen der Platten werden T-Zellpopulationen, Antigen-präsentierende Zellen (z.B. dendritische Zellen, Makrophagen, B-Zellen) und entsprechende Antigene (z.B. Peptide) eingebracht. Die Stimulation der T-Zellen erfolgt damit direkt auf der Platte. Zusätzlich werden negative Kontrollen (T-Zellpopulationen ohne Antigen und Antigen-präsentierende Zellen, T-Zellpopulationen mit Antigen-präsentierenden Zellen ohne Antigen) und positive Kontrollen (T-Zellpopulationen, Antigen-präsentierende Zellen, generell aktivierend wirkende Antigene oder unspezifische Direktstimulation durch Anti-CD3-Antikörper/Anti-CD28-Antikörper) angelegt. Es folgt eine 6-48h lange Inkubationszeit. Nach Antigenerkennung werden diese Antigen-spezifischen T-Zellen aktiviert und produzieren Zytokine, die von den am Boden der Platte haftenden Antikörpern gebunden werden. Die zellulären Bestandteile werden nach der Inkubationszeit in einem Waschschrift entfernt (Abb. 4a und 4b).

Mit einem zweiten Enzym-gekoppelten Anti-Zytokin-Antikörper werden die aus Antikörper und Zytokinen bestehenden Komplexe gebunden und nach Zugabe des Enzymsubstrates durch spezielle Farbreaktionen sichtbar gemacht (Abb. 4c und 4d). Derzeit finden zwei Färbemethoden Anwendung: das Alkali-Phosphatase-Verfahren, das die Spots in einem Blau-Ton zur Darstellung bringt und die horse-radish-Peroxidase (HRPO)-Technik, die bräunlich-rote Spots erzeugt.



**Abb. 4. a-d:** Das Elispot-Verfahren. Auf den Boden einer 96-Well-Mikrotiterplatte werden Zytokin-spezifische Antikörper aufgebracht. Sie binden die von Immunzellen nach Stimulation sezernierten Zytokine (a und b). Ein zweiter Enzym-markierter Anti-Zytokin-Antikörper wird zugesetzt und bindet an die Antikörper-Antigen-Komplexe. Nach Substratzugabe werden die Komplexe durch eine Farbreaktion sichtbar gemacht (c und d).

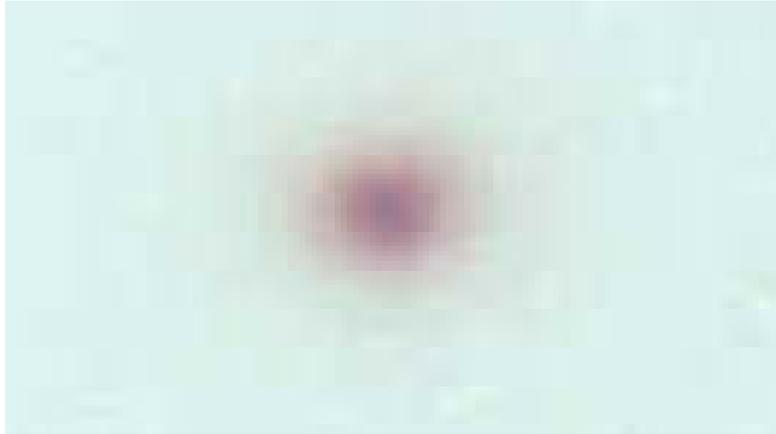
Jeder der so entstandenen, farbigen Areale (Spots) repräsentiert das Gebiet, innerhalb dessen eine T-Zelle mit Zytokinsekretion auf das Antigen reagiert hat. Jeder einzelne Spot entspricht somit dem Abdruck der ursprünglichen T-Zelle. Die auf der Oberfläche der Nitrocelluloseböden entstandenen Spots (Abb. 5) können dauerhaft in dieser Form aufbewahrt werden und stehen jederzeit, sowie beliebig oft für Auswertungen zur Verfügung.



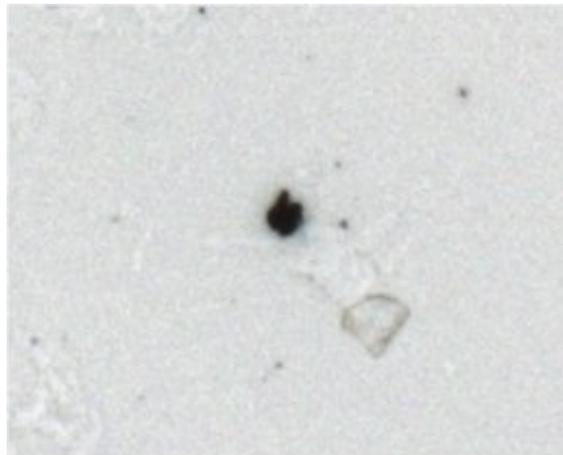
**Abb. 5:** Beispielhafte Aufnahme einer Vertiefung einer 96-Well-Mikrotiterplatte mit den darin durch den Elispot-Assay entstandenen Spots.

Wahre Spots besitzen einige typische Merkmale (Abb. 6), anhand derer sie von artifiziellen Farbpunkten (Abb. 7) oder Verunreinigungen abgegrenzt werden können:

- Echte Spots sind rund, gehen am Rand aber fließend in den Hintergrund über und haben somit keine scharfe Begrenzung
- Die Farbintensität ist im Zentrum eines echten Spots am höchsten und nimmt zum Rand hin kontinuierlich ab
- Falsche Spots sind häufig etwas kleiner, in ihrer Farbintensität homogen und weisen eine scharfe Randbegrenzung auf



**Abb. 6:** Beispielhafte Aufnahme eines typischen Spots. Die Zytokinkonzentration ist im Zentrum am höchsten und nimmt zur Peripherie hin fließend ab.



**Abb. 7:** Beispielhafte Aufnahme eines artifiziellen Spots. Der für echte Spots typische, fließende Übergang in den Hintergrund, sowie die zum Spotrand hin abnehmende Farbintensität fehlen. Der artifizielle Spot ist scharf begrenzt und homogen gefärbt.

### **2.1.3. Präparation der verwendeten ELISPOT-Proben**

Zur Herstellung der Human-IFN- $\gamma$ -ELISPOT-Assays wurden in die Vertiefungen der 96-Well-Multiscreen-Platten (Millipore, Burlington, USA) jeweils 100 $\mu$ l Anti-human-IFN- $\gamma$ -Antikörper von Mäusen (10  $\mu$ g/ml; clone 1-D1K; Mabtech, Schweden) gegeben. Die Platten wurden über Nacht bei 4°C aufbewahrt,

anschließend gewaschen und mit Humanserum 1 Stunde lang bei 37°C geblockt.

In die Wells wurden nach Isolierung mit Microbeads (Miltenyi Biotec, Auburn, USA) CD8<sup>+</sup>T-Zellen (1x10<sup>5</sup>/Vertiefung) mit autologen dendritischen Zellen (2x10<sup>4</sup>/Well) und Peptiden in Konzentration von 10µg/ml eingebracht. Einzelne Wells dienten als Negativkontrolle, manche von ihnen enthielten nur CD8<sup>+</sup>T-Zellen (und keine dendritischen Zellen), andere nur dendritische Zellen (und keine CD8<sup>+</sup>T-Zellen) oder nur Medium. Unspezifische Positivkontrollen stellten Wells dar, die CD8<sup>+</sup>T-Zellen, dendritische Zellen (2x10<sup>4</sup> /well) und 10µg/ml PHA-L (Lectin von Phaseolus vulgaris Leucoagglutinin) (Sigma) enthielten.

Nach einer Inkubationszeit von 20 Stunden bei 37°C wurden die Platten mit PBS/0.5% Tween gewaschen. Das am Boden der Platten von den Antikörpern gebundene IFN-γ wurde durch Zugabe von 100µl eines zweiten biotinylierten Anti-human-IFN-γ-Antikörpers gebunden (2 µg/ml; clone 7-B6-1; Mabtech, Schweden). Die Spots wurden durch Zugabe von 5µg/ml Avidin-Peroxidase-Complex (Vectastain Elite Kit, Vector), einen erneuten Waschvorgang und anschließende Zugabe von 3-Amino-9-Ethyl-Carbazol Substrat Lösung (AEC, Sigma, Taufkirchen, Deutschland) als bräunlich rote Farbpunkte sichtbar gemacht.

Die Maus-IFN-γ-Elispotplatten wurden von einer auswärtigen Arbeitsgruppe hergestellt (Ltg. Dr. A. Schmidt, apovia AG, 82152 Martinsried). Dabei wurde ein Maus-IFN-γ-Elispot-Set (BD Biosciences, San Diego, USA) verwendet. In die Vertiefungen einer 96-Well-Mikrotiterplatte wurden Anti-Maus-IFN-γ-Antikörper in einer Konzentration von 5µg/ml gegeben, mit RPMI 1640 (Biochrom, Berlin, Deutschland) gewaschen und mit RPMI 1640, das 10% fetales Kalb-Serum enthielt (PAA, Linz, Österreich), geblockt.

Milzzellen in Einzelzell-Suspension wurden in einer Konzentration von 2x10<sup>5</sup> /Vertiefung und in einer zweifach absteigenden Verdünnungsreihe hinzugegeben. Stimulierendes Peptid wurde bis zu einer Konzentration von 5µg/ml hinzugefügt. Nach einer Inkubationszeit von 36 bis 40 Stunden bei 37°C wurden die Platten zweimal mit destilliertem Wasser und dreimal mit PBS/0.05% Tween-20 gewaschen. Das gebundene IFN-γ wurde mit einem biotinylierten Anti-Maus-IFN-γ-Antikörper in einer Konzentration von 2µg/ml detektiert. Nach

Waschen der Platten mit PBS/0.05% Tween-20 wurden die Spots durch Zugabe von 5µg/ml Avidin-Peroxidase-Complex (Vectastain Elite Kit, Vector), erneuten Waschvorgang und anschließende Zugabe von 3-Amino-9-Ethyl-Carbazol Substrat Lösung (AEC, Sigma, Taufkirchen, Deutschland) als bräunlich rote Farbpunkte sichtbar gemacht.

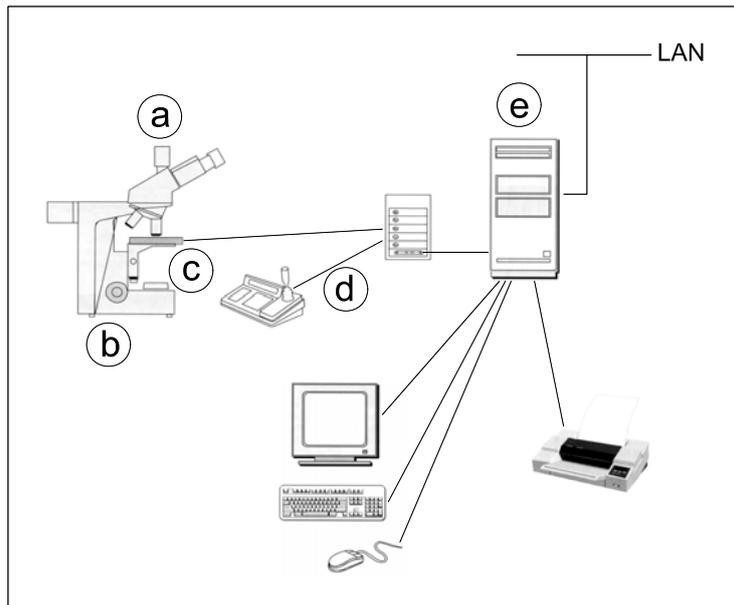
## **2.2. Automatisierte Bildanalyzesysteme**

Es wurden zwei Systeme mit insgesamt 3 verschiedenen Auflösungen zur automatisierten Auswertung von Elispot-Proben verwendet, das System KS Elispot und das System KS Elispot compact, entwickelt von Carl Zeiss Vision GmbH (Göttingen, Deutschland).

Die Systeme KS Elispot und KS Elispot compact bestehen aus den aufeinander abgestimmten Komponenten Mikroskop, Kamera, Motortisch mit Steuerung, Analysesoftware und Rechner. Die Spots werden im Mikroskop mit motorisiertem Tisch erfasst und über eine Farbkamera aufgenommen (Abb. 8).

Zur Auswertung können die auf einer Klebefolie aufgetragenen Nitrocelluloseböden aus den 96-Well-Mikrotiterplatten herausgelöst werden. Die Folie wird möglichst plan innerhalb einer Markierung auf dem motorisierten Tisch aufgelegt. Es stehen aber auch Halterahmen zur Verfügung, die auf dem motorisierten Tisch aufgebracht werden und in die die komplette 96-Well-Mikrotiterplatte eingelegt werden kann. In dieser Arbeit wurden beide Anordnungen verwendet.

Nach Digitalisierung der aufgenommenen Well-Bilder mit einer 24-Bit Farbauflösung, werden sie von der KS Elispot Software ausgewertet. Die Ergebnisse erscheinen auf dem Bildschirm, dabei können Anzahl, Größe, Fläche und Intensität der Spots bestimmt und die Auswertungen in einem Reportformular mit Übersichtsbildern der einzelnen Wells und der darin erkannten Spots gespeichert werden.



**Abb. 8:** Schematische Darstellung der Komponenten der Systeme KS Elispot und KS Elispot compact. a=Kamera, b=Mikroskop, c=Motortisch, d=Steuereinheit des Motortisches, e= Systemrechner mit Analysesoftware.

### 2.2.1. KS Elispot

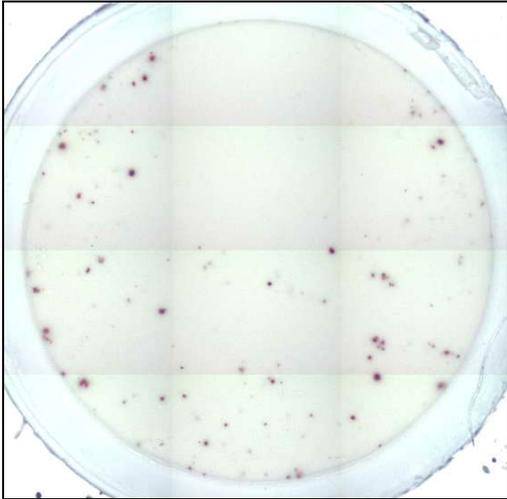
Das KS Elispot System (Abb. 9) ist mit einem Auflichtmikroskop, dem Zeiss Axioplan 2 imaging, mit Z-Schrittmotor und Getriebe, Epiplan Neofluar 5x / 0,15 Objektiv und Zoomadapter ausgestattet. Es verfügt über einen motorisierten Tisch (Märzhäuser Elispot Motortisch 4x5 Zoll), der über eine 3-Achsen-Steuereinheit kontrolliert wird. Über Standardschnittstellen, CAN-Bus und RS 232, arbeiten der eingebaute Mikroprozessor und ein Steuerprogramm mit dem Systemrechner (INTEL P4, 256 MB RAM) zusammen, so dass eine Prozessautomatisierung ermöglicht wird. Die Anpassung des auszuwertenden Plattenbereiches erfolgt manuell mittels Joystick. Pro Well werden 12 Einzelbilder erstellt (Abb. 10a, b) und von einer 3 Chip CCD (Charge-Coupled-Device)-Farbkamera (Sony DXC 950) aufgenommen.

Nach Digitalisierung der aufgenommenen Well-Bilder mit einer 24-Bit Farbauflösung, werden sie von der KS Elispot Software (Version 4. 3) ausgewertet. Die erreichte Pixelgröße bzw. die Bildpunktauflösung für ein

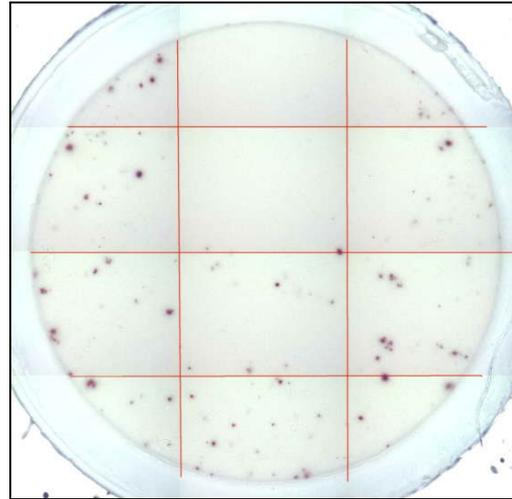
760x580 Pixel Bild eines Wells beträgt  $2.6\mu\text{m}$ . Ein Pixel (Bildpunkt) entspricht bei der gerasterten, digitalisierten Bilddarstellung dem kleinsten, darstellbaren Element des Bildes. Die digitalisierte Bildverarbeitung basiert auf der Zerlegung des Bildes nach Spalten und Reihen in einzelne Pixel.



**Abb. 9:** Das System KS Elispot zur automatisierten, bildanalytischen Auswertung von Elispot-Proben mit den Systemkomponenten Kamera (a), Mikroskop (b), Steuereinheit des Motortisches (c) und Motortisch (d) mit aufliegender, aus der Elispotplatte herausgelöster Klebefolie.



**Abb. 10a**



**Abb. 10b**

**Abb. 10a, b:** Beispielhafte Aufnahme eines Wells (a), erstellt am System KS Elispot. Das System nimmt 12 Einzelbilder pro Well auf, verdeutlicht durch Rasterung (b).

### **2.2.2. KS Elispot compact**

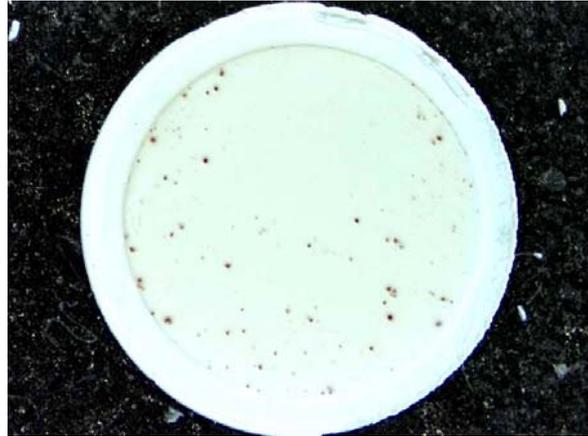
Die Elispot-Proben werden bei diesem System (Abb. 11) über ein Stereomikroskop, das Zeiss Stemi 2000 C, mit Auflichtbeleuchtung und motorisiertem Scanning-Tisch (Märzhäuser Elispot Motortisch 4x5 Zoll) vermessen. Der Motortisch wird über eine 2-Achsen-Steuereinheit kontrolliert, das auszuwertende Plattenareal wird mittels Joystick markiert. Eine 3 Chip CCD-Kamera (Hitachi HV C20A) nimmt die Zytokin-Spots auf. Nach Digitalisierung der aufgenommenen Well-Bilder mit einer 24-Bit Farbauflösung, werden sie von der KS Elispot Software (Version 4. 3) ausgewertet.

Die Auswertungen wurden bei 2 Mikroskopvergrößerungen durchgeführt, dem 0.65-Zoom, sowie dem 1.25-Zoom. Bei Verwendung des 0.65-Zooms wird pro Well ein einziges Bild aufgenommen (Abb. 12). Mit dieser Vergrößerung können Spots ab einem Durchmesser von 30 $\mu$ m erkannt werden. Für ein 760x580 Pixel Bild eines Wells resultiert eine Pixelgröße von 12 $\mu$ m. Bei Wahl des 1.25-Zooms werden von jedem Well vier Einzelaufnahmen (Abb. 13a, b) erstellt, die zu einem Gesamtbild des Wells zusammengesetzt werden. Die Verwendung von mehreren Bildern pro Well mit einer höheren Primärvergrößerung führt dazu,

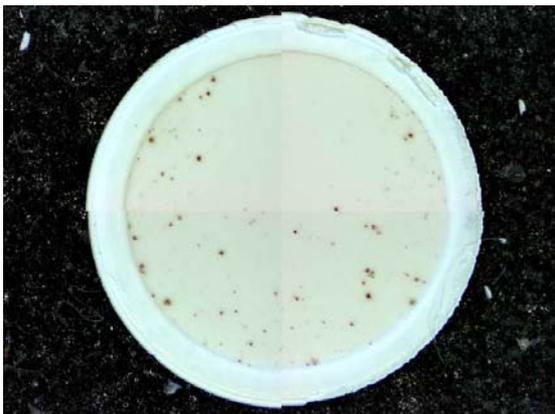
dass Spots mit einem Durchmesser ab  $20\mu\text{m}$  erfasst werden können. Die Bildpunktauflösung bei einem  $760 \times 580$  Pixel Bild liegt bei  $6\mu\text{m}$ .



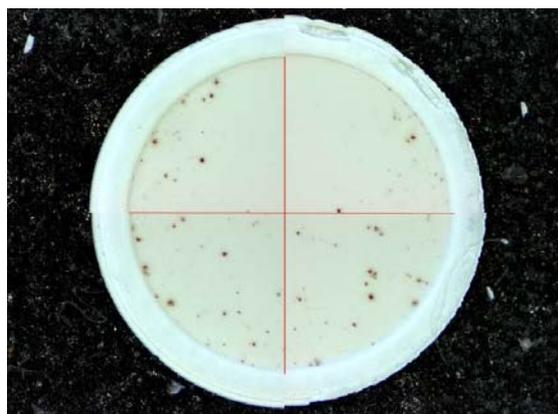
**Abb. 11:** Das System KS Elispot compact zur automatisierten Auswertung von Elispot-Proben mit den Systemkomponenten Kamera (a), Mikroskop (b), Motortisch mit im speziellen Halterahmen aufliegender Elispot-Platte (c).



**Abb. 12:** Beispielhafte Aufnahme desselben Wells wie Abb. 10, erstellt am System KS Elispot compact unter Verwendung der 0.65-Zoom-Mikroskopvergrößerung, die 1 Bild pro Well anfertigt.



**Abb. 13a**



**Abb. 13b**

**Abb. 13a, b:** Beispielhafte Aufnahme desselben Wells (a) wie Abb. 10 und 12, erstellt am System KS Elispot compact unter Verwendung der 1.25-Zoom-Mikroskopvergrößerung. Es werden 4 Einzelbilder pro Well aufgenommen und zum Gesamtbild zusammengefügt (b).

| <b>Systemkonfiguration</b>                   |   |
|--|---|
| <b>System KS Elispot</b>                     | <b>System KS Elispot compact</b>                                  |
| Mikroskop Axioplan 2                         | Mikroskop Stemi 2000C   |
| Autofokuseinrichtung                         | -   |
| Motortisch mit<br>Steuereinheit und Joystick | Motortisch mit<br>Steuereinheit und Joystick                      |
| 3 Chip CCD Farbkamera                        | 3 Chip CCD Farbkamera   |
| KS Elispot Software Version 4.3              | KS Elispot Software Version 4.3                                   |
| Systemrechner                                | Systemrechner   |
| 12 Aufnahmen pro Well                        | 0.65-Zoom: 1 Aufnahme pro Well<br>1.25-Zoom: 4 Aufnahmen pro Well |

**Tab. 1:** Gegenüberstellung der Systemkomponenten des KS Elispot und des KS Elispot compact.

### 2.2.3. KS Elispot Software

Die KS Elispot Software verfügt über einen Spot-Erkennungs- und Lernalgorithmus. Der Spoterkennungsalgorithmus dient zur Erkennung von Spots und zur Abgrenzung dieser „wahren“ Spots gegen artifizielle Farbpunkte oder Verunreinigungen. Wichtig für die Unterscheidung sind die typischen Merkmale eines Spots. Als Spots werden runde Farbareale mit am Rand fließendem Übergang zum Hintergrund und einer vom Zentrum zur Peripherie hin kontinuierlich abnehmenden Farbintensität bezeichnet. Eine scharfe Begrenzung zum Hintergrund und eine gleichmäßige Färbung, wie sie oft bei artifiziellen Farbpunkten, so genannten „falschen“ Spots vorkommt, fehlen bei „wahren“ Spots. Unter Einbezug dieser Kriterien wurden 6 Systemparameter für die Spoterkennung definiert (Tab. 2)

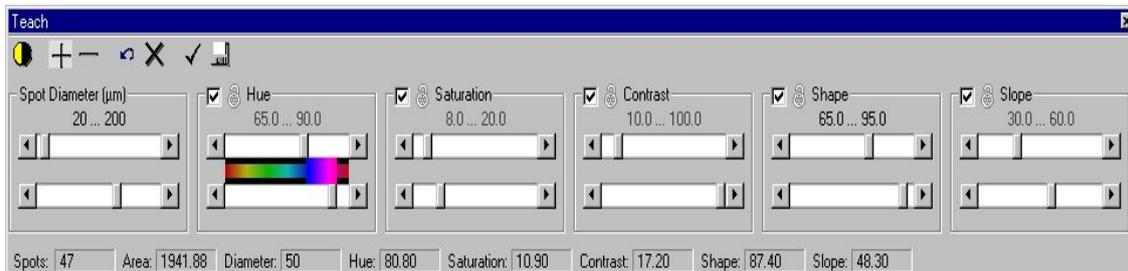
| Parameter                             | Parameterbeschreibung                        |
|---------------------------------------|--|
| Spotdurchmesser (Diameter)            | Größe des Spots                              |
| Farbwert (Hue)                        | Farbe des Spots                              |
| Farbsättigung (Saturation)            | Farbtiefe und Farbintensität des Spots       |
| Kontrast (Contrast)                   | Spotkontrast zum Untergrund                  |
| Form (Shape)                          | Runde Form des Spots                         |
| Randsteigung/ Kantensteilheit (Slope) | Fließende Farbverblässung zur Peripherie hin |

**Tab. 2:** Zusammenfassung der 6 Parameter eines Spots, die in den Erkennungsalgorithmus der KS Elispot Software implementiert wurden. Sie können im Rahmen der Erstellung eines „settings“, der Spotparameterkonstellation, individuell angepasst und abgespeichert werden. Die Reproduzierbarkeit der Messungen wird dadurch gesteigert.

In Abhängigkeit von zahlreichen Faktoren, so z. B. von verwendeten Antikörpern, Antikörperkonzentrationen, von verschiedenen Zytokinen oder Mikrotiterplatten, können die Spots innerhalb bestimmter Grenzen in ihrer Formation variieren. Vor Beginn einer Messung wird durch den Untersucher die Einstellung der Systemparameter für die Spoterkennung optimiert. Das Spot-Anlern-Verfahren, ein speziell entwickelter Lernalgorithmus zur Spoterkennung, dient der Anpassung der einzelnen Parameter an die gewünschten Spots im Bild.

Der Untersucher nimmt diese Parametereinstellung durch Auswahl der „wahren“, noch nicht erkannten Spots mit Cursor und linker Maustaste vor. Bei Auswahl eines Spots werden alle anderen, diese Kriterien ebenfalls erfüllenden Spots automatisch miteinbezogen und die Systemparameter entsprechend neu angepasst (Abb. 14). Ebenso kann der Untersucher bestimmte Spots und ihre Spotparameter durch Auswahl ausschließen. Die ausgewählten Parameter können in einem „setting“ abgespeichert werden, so dass sie bei künftigen Messungen wieder aufgerufen und verwendet werden können. Die

Spotauswahlkriterien müssen somit nicht vor jeder Messung neu bestimmt werden. Die Reproduzierbarkeit der Messungen wird durch dieses Verfahren deutlich erhöht.



**Abb. 14:** Die Systemparameter zur Spoterkennung (Spot Diameter, Hue, Saturation, Contrast, Shape, Slope) werden im Teach-Modus durch Auswahl der Spots mit dem Cursor angepasst. (Vgl. Tab. 2).

Bei den beiden untersuchten Reader-Systemen KS Elispot und KS Elispot compact ist die Übertragung der an einem System eingestellten Spotauswahlparameter auf das andere aus technischen Gründen nicht sinnvoll. Aufgrund der unterschiedlichen Beleuchtungsverfahren der Mikroskope, der abweichenden Entfernungen zwischen Mikroskop und Motortisch, und weiterer differenter Einstellungen an beiden Geräten, erwiesen sich die vom KS Elispot System übernommenen Einstellparameter am KS Elispot compact als unbrauchbar und ermöglichten keine zuverlässige Spoterkennung, so dass an jedem der beiden Systeme eine eigene Anpassung der Spotparameter vorzunehmen war.

## **2.3. Auswertungsprogramm der Elispotproben**

### **2.3.1. Einstellung der Systemparameter zur Spoterkennung**

Um die in der täglichen Praxis herrschenden Bedingungen bestmöglich zu imitieren, wurde in dieser Arbeit vor Beginn der Auswertungen für jedes System die individuell optimale Spotparameterkonstellation, das Setting, erstellt. Eines für das System KS Elispot und zwei am System KS Elispot compact, jeweils für die 0.65-Zoom- und 1.25-Zoom-Einstellung. Die Anpassung der Spotparameter erfolgte unter Verwendung des Spotanlernverfahrens. Spots deren Parameter miteinbezogen werden sollten, wurden ausgewählt, artifizielle Farbpunkte und deren Parameter ausgeschlossen, bis schließlich die Parametereinstellung feststand, die die Erkennung „wahrer“ Spots möglichst zuverlässig ermöglichte.

Um die drei Parametereinstellungen zu prüfen, wurden von einzelnen Wells mit den darin als erkannt markierten Spots am KS Elispot, am KS Elispot compact 0.65-Zoom-Einstellung und am KS Elispot compact 1.25-Zoom-Einstellung Bilder aufgenommen und ausgedruckt. Anhand der drei Aufnahmen eines Wells wurde visuell geprüft und bestätigt, dass die drei Parametereinstellungen soweit an die vorliegenden Spots angepasst wurden, dass sie dieselben „wahren“ Spots erkannten. Die Systemeinstellungen für die Spoterkennung wurden abgespeichert und ohne Änderung für alle nachfolgenden Messungen verwendet.

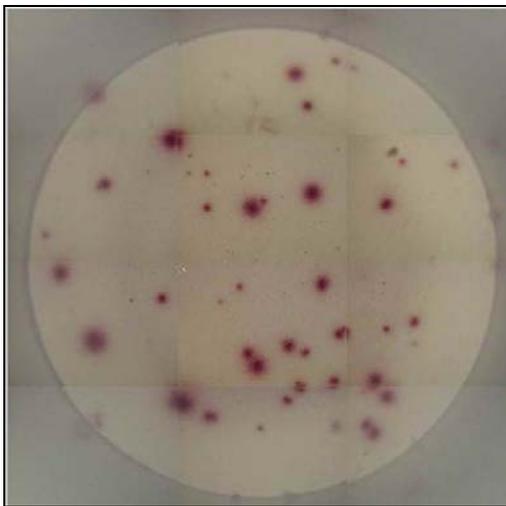
### **2.3.2. Auswertung der Elispotproben**

Am System KS Elispot und am System KS Elispot compact mit der 0.65-Zoom-Einstellung sowie mit der 1.25-Zoom-Einstellung wurden die folgenden Elispot-Proben ausgewertet:

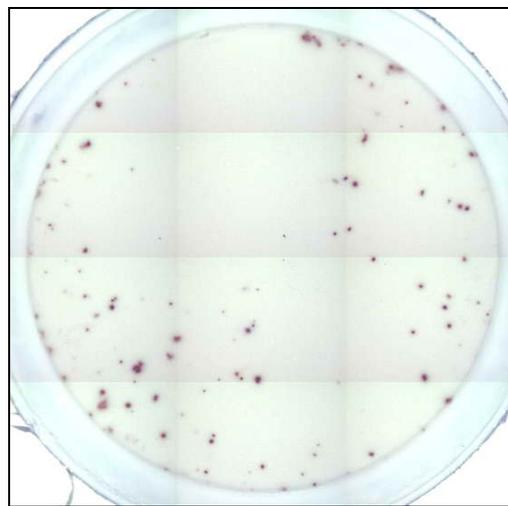
- 4 Maus-IFN- $\gamma$ -Mikrotiterplatten mit insgesamt 361 ausgewerteten Wells. Die Spotanzahl der einzelnen Wells reichte von 0 bis 283 Spots pro Well.
- 3 Human-IFN- $\gamma$ -Mikrotiterplatten mit insgesamt 226 ausgewerteten Wells. Die Spotanzahl der einzelnen Wells reichte von 0 bis 906 Spots pro Well.

Zur Auswertung der Maus-Elispot-Proben wurden auf dem motorisierten Tisch der Lesesysteme aufzubringende Halterahmen verwendet, in die die komplette 96-Well-Mikrotiterplatte eingelegt und ausgewertet werden konnte.

Für die Auswertung der Human-Elispot-Proben wurden die auf einer Klebefolie aufgetragenen Nitrocelluloseböden aus den 96-Well-Mikrotiterplatten herausgelöst und auf dem motorisierten Tisch der Systeme aufgelegt. Die untenstehenden Aufnahmen (Abb. 15a, b) zeigen beispielhaft Maus- und Humanspots aus den verwendeten Elispot-Platten.



**Abb.15a**



**Abb. 15b**

**Abb. 15a, b:** Die Abbildung 15a zeigt beispielhaft Maus-IFN- $\gamma$ -Spots im Well einer Mikrotiterplatte. In der Abbildung 15b sind Human-IFN- $\gamma$ -Spots dargestellt. Beide Aufnahmen wurden am System KS Elispot erstellt. Zur Auswertung der Human-Elispot-Proben wurden die abgelösten Nitrocelluloseböden der Mikrotiterplatten auf der Platte des Motortisches aufgebracht, während die Maus-Elispot-Platten in spezielle Halterahmen eingelegt wurden.

Für die zwei Systeme KS Elispot und KS Elispot compact (0.65-Zoom, 1.25-Zoom) wurden in Anlehnung an die vom jeweiligen Auswertungssystem pro Well aufgenommene Bilderanzahl die folgenden Bezeichnungen gewählt:

System KS Elispot:

12 Bildaufnahmen pro Well → Reader 12 oder R12

System KS Elispot compact 0.65-Zoom:

1 Bildaufnahme pro Well → Reader 1 oder R1

System KS Elispot compact 1.25-Zoom:

4 Bildaufnahmen pro Well → Reader 4 oder R4

Die Messdaten vom KS Elispot compact wurden mit den vom System KS Elispot ermittelten Werten verglichen, demnach also R1 mit R12 und R4 mit R12.

Die Auswertung der Elispot-Proben erfolgte in 2 getrennten Messungen, der Messanordnung 1 und der Messanordnung 2, mit jeweils unterschiedlichen Messbedingungen:

#### Messanordnung 1

Unter den Messbedingungen der Messanordnung 1 wurden Spots des gesamten Größenbereiches 20-300µm in die Messung einbezogen. Jede der Maus- und Human-Mikrotiterplatten wurde an den verglichenen Systemen einmal ausgewertet. Damit wurden im Rahmen der durchmesserunabhängigen Messanordnung 1 an den verglichenen Systemen R12, R4 und R1 jeweils 587 Wells ausgewertet.

#### Messanordnung 2

Bei dem Systemvergleich interessierte, ob und wie die Größe der einzelnen Spots auf die Ergebnisse der beiden Geräte Einfluss nimmt.

In dieser Messung wurden die Messbedingungen daher so verändert, dass nur Spots bestimmter Größen bzw. bestimmter Durchmesser in die Auswertung einbezogen wurden. Die Spoterkennungsparameter am System können durch

Eingabe des minimalen und maximalen Spotdurchmessers so eingestellt werden, dass nur Spots des gewünschten Spotgrößenbereiches in die Messung eingehen.

Für die Maus-Mikrotiterplatten wurden zwei Durchmesserbereiche festgelegt:

- 20-100µm
- 100-300µm

Für die Human-Mikrotiterplatten wurden drei Spotdurchmesserbereiche gewählt:

- 20-50µm
- 50-100µm
- 100-300µm

Die Humanplatten enthielten im Vergleich zu den Maus-Elislot-Proben mehr kleine Spots. Um die Unterschiede der Messsysteme in den kleineren Größenbereichen klarer herauszuarbeiten, wurden für Humanspots daher drei Größenkategorien gewählt.

Im Rahmen dieser durchmesserabhängigen Messung wurden alle Maus-Mikrotiterplatten jeweils zwei Mal (zur Erfassung von Spots der Durchmesser 20-50µm und 100-300µm) an den Systemen R12, R4 und R1 gelesen. Damit wurden an jedem der verglichenen Systeme 722 Wells ausgewertet.

Alle Human-Mikrotiterplatten wurden drei Mal (zur Erfassung von Spots der Durchmesser 20-50µm, 50-100µm und 100-300µm) an den Systemen R12, R4 und R1 gelesen. An den verglichenen Systemen wurden somit jeweils 678 Wells ausgewertet.

Insgesamt wurden im Rahmen der durchmesserabhängigen Messanordnung 2 an jedem der verglichenen Systeme R12, R4 und R1 jeweils 1400 Wells ausgewertet.

Darüber hinaus wurden weitere Auswertungen von Wells durchgeführt. Sie dienten zur Erfassung der von den Systemen benötigten Auswertungszeit, zur Bestimmung der Variabilität innerhalb mehrerer Auswertungen und zur Untersuchung der Zuverlässigkeit (reliability) der Spoterkennung der beiden Systeme (Siehe 2. 5).

## **2.4. Bewertung und Darstellung von Messergebnissen**

### **2.4.1. Qualitätsparameter zur Bewertung von Messergebnissen**

In dieser Arbeit wurden die zwei Bildanalyseysteme KS Elispot und KS Elispot compact anhand der von ihnen ermittelten Messdaten miteinander verglichen, um Aussagen darüber treffen zu können, wie gut sie in ihren Ergebnissen übereinstimmen.

Zur Beurteilung von Messergebnissen, die ein Messgerät liefert, dienen im Allgemeinen die drei Begriffe:

- Genauigkeit (accuracy)
- Richtigkeit (trueness, accuracy of the mean)
- Präzision (precision)

Richtigkeit und Präzision sind Kriterien der Genauigkeit.

Die Genauigkeit beschreibt das Ausmaß der Übereinstimmung des einzelnen Messergebnisses mit dem wahren Wert der Messgröße, sie definiert wie nahe das Ergebnis am wahren Wert liegt.

Nach DIN 55350 (Deutsches Institut für Normung (DIN), 1997) ist die Genauigkeit definiert als qualitative Bezeichnung für das Ausmaß der Annäherung von Ermittlungsergebnissen an den Bezugswert, wobei dieser je nach Festlegung oder Vereinbarung der wahre, der richtige oder der Erwartungswert sein kann.

Der wahre Wert gilt naturgemäß als Idealwert, den man in einer perfekten Messung erhalten würde und ist demnach unbestimmt. Der wahre Wert einer Messgröße wird nach DIN 55350 definiert als der tatsächliche Wert einer Messgröße unter den, bei der Ermittlung herrschenden Bedingungen.

Den Mittelwert oder auch das arithmetische Mittel erhält man durch Addition der Einzeldaten und Division dieses Wertes durch die Datenanzahl. Im Gegensatz zum wahren Wert ist der Mittelwert mit einem Fehler behaftet. Die Abweichung zwischen dem wahren Wert und Mittelwert ist der bias (siehe unten: systematischer Fehler).

Für eine Anzahl  $n$  von Einzeldaten  $x_i$  ergibt sich für den Mittelwert  $\bar{x}$  :

$$\bar{x} = \sum \frac{x_i}{n}$$

Die Standardabweichung  $[s]$  oder das Streumaß gibt an, wie sich die einzelnen Messdaten um den Mittelwert verteilen:

$$s = \pm \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Die Richtigkeit ist ein Maß für die Abweichung zwischen dem Mittelwert einzelner Messergebnisse und dem wahren Wert der Messgröße. Sie beschreibt wie gut Mittelwert und wahrer Wert übereinstimmen.

Nach DIN 55350 ist die Richtigkeit definiert als qualitative Bezeichnung für das Ausmaß der Annäherung des Erwartungswertes des Ermittlungsergebnisses an den Bezugswert, wobei dieser je nach Festlegung oder Vereinbarung der wahre oder der richtige Wert sein kann.

Die Präzision ist ein Maß für die Übereinstimmung unabhängiger Messergebnisse unter festen Bedingungen. Sie gibt die Reproduzierbarkeit oder Wiederholbarkeit eines Messergebnisses an.

Nach DIN 55350 ist die Präzision definiert als qualitative Bezeichnung für das Ausmaß der gegenseitigen Annäherung voneinander unabhängiger Ermittlungsergebnisse bei mehrfacher Anwendung eines festgelegten Ermittlungsverfahrens unter vorgegebenen Bedingungen.

Die Genauigkeit vereinbart mit den Begriffen Richtigkeit und Präzision auch die Effekte zweier Fehlerarten, die bei jedem Messverfahren auftreten. Es sind die Effekte des systematischen und des statistischen Fehlers.

### Systematischer Fehler

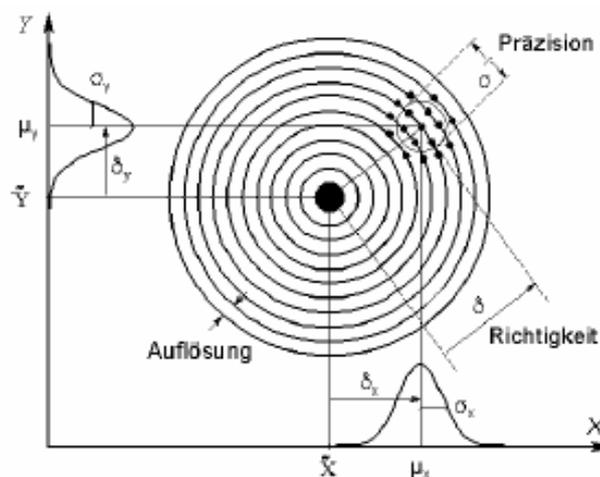
Der systematische Fehler (bias,  $b$ ) bestimmt die Richtigkeit des Messverfahrens. Er ist definiert als die Differenz zwischen dem Mittelwert  $\mu$  und dem wahren Wert  $x$ , systematischer Fehler  $b = x - \mu$ .

Systematische Fehler entstehen durch nicht optimierte, fehlerhafte Messverfahren und Messgeräte oder durch störende Einflüsse.

### Statistischer (zufälliger) Fehler

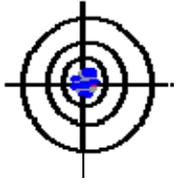
Der statistische Fehler bestimmt die Präzision des Messverfahrens. Statistische Fehler sind nicht beherrschbar und lassen sich nicht vollständig eliminieren. Die Präzision kann als Standardabweichung quantifiziert werden. (Kreyszig, 1968), (Sachs, 1978), (Koch, 1987), (Harms, 1997).

Die Begriffe Richtigkeit und Präzision sowie die Fehlerarten systematischer und statistischer Fehler lassen sich anhand des Vergleichs mit einer Schussscheibe demonstrieren (Abb. 16).

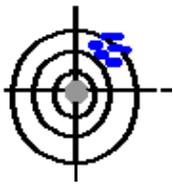


**Abb. 16:** Präzision und Richtigkeit als Kriterien der Genauigkeit, dargestellt am Beispiel einer Zielscheibe.

Die verschiedenen Fehlerkombinationen können am Beispiel einer Schießscheibe demonstriert werden:



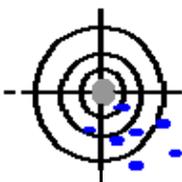
**Abb. 17:** Präzise und richtig, kein bias. Die Treffer liegen im Inneren der Zielscheibe dicht aneinander. Die Präzision ist hoch, der statistische Fehler ist klein. Kein systematischer Fehler.



**Abb. 18:** Präzise und falsch, enthält bias. Die Treffer liegen dicht aneinander, jedoch nicht im Mittelpunkt. Die Präzision ist hoch, der systematische Fehler jedoch groß, das Ergebnis ist falsch.



**Abb. 19:** Nicht präzise und richtig, kein bias. Die Treffer liegen weit voneinander verstreut, die Präzision ist niedrig, der statistische Fehler hoch. Da das Ergebnis im Mittelwert richtig ist, ist der systematische Fehler null.



**Abb. 20:** Nicht präzise und falsch, enthält bias. Die Treffer liegen weit voneinander verstreut und sind vom Mittelpunkt nach rechts verschoben. Die Präzision ist niedrig, der statistische und systematische Fehler sind hoch.

Um die Systeme KS Elispot und KS Elispot compact bezüglich der Begriffe Genauigkeit, Richtigkeit und Präzision charakterisieren zu können, wurde geprüft welches graphisch-statistische Verfahren am besten zur Visualisierung und Quantifizierung der Unterschiede zwischen den zwei Messgeräten geeignet ist.

#### **2.4.2. Darstellung von Ergebnissen in Methoden-vergleichenden Studien**

Es existieren mehrere Verfahren, die in Methoden-vergleichenden Untersuchungen angewendet werden können:

- Korrelations- bzw. Regressionsanalyse
- Passing-Pablok-Regression
- Varianzanalytische Methoden (Intraklassen-Korrelationskoeffizient)
- Mountain-Plots
- Bland-Altman-Plots

##### **Korrelations-bzw. Regressionsanalyse**

Dieses Verfahren sieht die Erstellung eines Streu-oder Korrelationsdiagramms (eng. Scatterplot) mit der anschließenden Durchführung einer Korrelations-und Regressionsanalyse vor. In einem Scatterplot werden im Koordinatensystem auf der Abszisse die Ergebnisse der einen Methode, auf der Ordinate die Ergebnisse der anderen Methode aufgetragen.

Mittels Korrelationsanalyse und Berechnung des Korrelationskoeffizienten kann ermittelt werden, ob zwischen den Variablen  $x$  und  $y$  ein linearer stochastischer Zusammenhang besteht und abgeleitet werden, wie groß die Übereinstimmung der verglichenen Methoden ist. Im Falle, dass beide Methoden exakt übereinstimmen, kämen alle Punkte auf der Geraden  $y=x$  (Winkelhalbierende im Koordinatensystem) zu liegen, der Korrelationskoeffizient hätte den Wert 1. Dies kommt in der Realität schon aufgrund von Messfehlern praktisch nicht vor, so dass Korrelationskoeffizienten ab 0.8 als befriedigend angesehen werden. Mittels Regressionsanalyse wird ermittelt welcher Zusammenhang zwischen den Variablen  $y$  und  $x$  besteht und ob sich  $y$  aus  $x$  schätzen lässt.

Sind beide Variablen der Regressionsanalyse fehlerbehaftet, muss festgelegt werden, welche Variable als abhängig und welche als unabhängig anzusehen ist. Ist diese willkürliche Wahl nicht vertretbar, führt dies zum relativ komplexen Problem der linearen Strukturgleichungen (structural relationship). In jedem Fall wird vorausgesetzt, dass beide Variablen unabhängig voneinander mit konstanter Varianz annähernd normal verteilt sind.

### **Passing-Pablok-Regression**

Beim wenig bekannten linearen Regressionverfahren nach Passing und Pablok wird der Regressionskoeffizient durch die Ranginformationen der Steigungen zwischen den Messwertpaaren geschätzt. Es kann auch unter schwachen Voraussetzungen angewandt werden und wird daher zu den nicht-parametrischen Verfahren gerechnet.

### **Intraklassen-Korrelationskoeffizient**

Varianzanalytische Verfahren ziehen als Maßzahl zur Beurteilung von Messwertpaaren die Varianzen zwischen den Messwerten und denen innerhalb der Messwerte heran. Die gebräuchlichste ist der Intraklassen-Korrelationskoeffizient. Wie bei der Korrelation müssen dabei gewisse Verteilungsannahmen erfüllt sein.

### **Äquivalenztests**

Üblicherweise werden die Methoden mittels der üblichen Signifikanztests (z.B. gepaarter t-Test, Rangtest) verglichen. Bei genügend großen Stichprobenumfängen (Power) erhält man allerdings nur das Resultat, dass die Methoden verschieden sind, was i. A. aber schon von vornherein bekannt ist, andernfalls erübrigen sich weitere Analysen. Als Testverfahren kämen allenfalls Äquivalenz-Tests infrage, bei denen Null- und Alternativhypothese vertauscht sind. Dazu muss im Voraus ein Intervall definiert werden, innerhalb dessen die (mittleren) Differenzen der Messwerte als äquivalent, d.h. akzeptabel übereinstimmend angesehen werden können.

### **Mountain-Plots**

Beim sog. Mountain-Plot (folded empirical cumulative distribution plot) (Krouwer and Monti, 1995) wird die kumulative Verteilung der Werte-Differenzen gegen die Differenzen selbst aufgetragen, d.h. man kann ablesen, welche Differenzen wie oft auftreten. Mountain-Plots sind als ergänzende Darstellungen zu anderen Methoden gedacht.

### **Bland-Altman-Plots**

Alle bisher erwähnten Verfahren haben gewisse Mängel. Die Nachteile der Regressions- und Korrelationsverfahren bei ihrer Anwendung in Methodenvergleichenden Untersuchungen werden bei Bland und Altman diskutiert (Bland and Altman, 1986), (Bland and Altman, 2003). Sie bemängeln in ihren Ausführungen die falsche Interpretation des Korrelationskoeffizienten. Dieser würde den Zusammenhang zweier Variablen, jedoch nicht deren Übereinstimmung messen, die in Methodenvergleichenden Untersuchungen von Interesse ist. Ein hoher Korrelationskoeffizient sei damit nicht gleichbedeutend mit einer guten Übereinstimmung der verglichenen Methoden. Weiterhin führen sie auf, dass bei den anderen Verfahren eventuelle Abhängigkeiten des Fehlers hinsichtlich Umfang und Richtung von der Größe der Messwerte nicht oder nur wenig berücksichtigt werden.

Bland und Altman entwickelten daher die so genannten Bland-Altman-Plots. Sie gelten als anerkanntes Verfahren zur Ergebnisdarstellung beim Vergleich zweier Messmethoden in medizinischen Studien. Ihre Bedeutung und Verbreitung nehmen stetig zu, viele Wissenschaftler tragen mit ihren Erfahrungsberichten und Anregungen zur ihrer ständigen Optimierung bei (Petersen et al., 1997), (Mantha et al., 2000), (Dewitte et al., 2002).

Bei der Datenanalyse mittels Bland-Altman-Plots werden auf der Abszisse die Mittelwerte der Ergebnisse, die von den zu vergleichenden Messmethoden ermittelt wurden, gegen die absolute Differenz der Wertepaare auf der Ordinate aufgetragen. Übereinstimmende Werte (Differenz=0) liegen dabei auf der Abszissenachse. Wenn die Fehler proportional zu den Werten anwachsen, können alternativ auch die Differenzen der Logarithmen gegen die Mittelwerte oder, äquivalent, deren Verhältnis aufgetragen werden.

Zur Visualisierung der Richtigkeit wird der Mittelwert der Differenzen als Parallele zur x-Achse eingezeichnet. Weiterhin ist es üblich, unter der Annahme einer Normalverteilung der Differenzen, die Grenzen zu bestimmen, innerhalb deren 95% aller Differenzen zu erwarten sind. Sie werden aus dem Mittelwert aller Differenzen und ihrer Standardabweichung berechnet. Für die obere bzw. untere Grenze gilt: Mittelwert der Differenzen + bzw. - 1.96\*Standardabweichung der Differenzen. Bland und Altman bezeichnen diese Grenzen als die „95%-limits of agreement“. Diese „Übereinstimmungsgrenzen“, auch als die 95%-Konfidenzgrenzen oder die 95%-Vertrauensintervalle bezeichnet, schließen den Bereich ein, in dem 95% aller Differenzen erwartet werden können und dienen zur Beurteilung der Genauigkeit. (Bland and Altman, 1995), (Bland and Altman, 1999).

Sind die Differenzen unabhängig von den auf der x-Achse aufgezeichneten Werten, werden die Grenzen parallel zur x-Achse eingezeichnet. Für den Fall, dass die Differenzen mit den Werten auf der x-Achse zusammenhängen, so dass Mittelwerte und Standardabweichung der Differenzen mit steigender Anzahl zunehmen, können diese Grenzen nicht mehr parallel zur x-Achse eingetragen werden.

Bland und Altman empfehlen in diesem Fall anstatt den Mittelwert der Differenzen als Parallele zur x-Achse einzuzichnen, die Regressionsgerade zu verwenden und die Konfidenzgrenzen V-förmig um die x-Achse anzuordnen. Für solche nicht-parallele Grenzen der 95%-Vertrauensintervalle geben sie entsprechende Formeln an. Andere Untersucher favorisieren die prozentuale Darstellung der y-Achse, da sie eine bessere Visualisierung der Ergebnisse und das Einzeichnen der Konfidenzgrenzen parallel zur Abszisse ermöglicht. (Pollock et al., 1992).

Für den Fall, dass ein "Gold-Standard" vorliegt, d.h. eine Messwertreihe bekannt (richtig) ist, taucht die Frage auf, ob dann besser die Abweichungen gegen die bekannten (wahren) Werte aufgetragen werden sollten, anstatt gegen die Mittelwerte. Bland und Altman haben nachgewiesen, dass in diesem Fall eine inhärente Korrelation zwischen Abszissen- und Ordinatenwerten vorliegt und somit den Zusammenhang verfälscht.

Als Darstellungsverfahren für die Ergebnisse des Vergleichs der zwei Lesesysteme bzw. der drei unterschiedlichen Messanordnungen wurden in dieser Arbeit die Bland-Altman-Plots gewählt. Sie geben die Qualitätsparameter zur Beurteilung von Messergebnissen sehr gut wider und zeigen das Ausmaß der Übereinstimmung zweier Messmethoden.

### **2.4.3. Darstellung der Ergebnisse mittels Bland-Altman-Plots**

In dieser Arbeit wurden in den Bland-Altman-Plots auf der Abszisse die Mittelwerte der von beiden Geräten pro Well gezählten Spotanzahl aufgetragen. Beim Vergleich von KS Elispot gegen KS Elispot compact 0.65-Zoom:  $[(R12+R1)/2]$ . Beim Vergleich von KS Elispot gegen KS Elispot compact 1.25-Zoom:  $[(R12+R4)/2]$ . Auf der Ordinate wurden die absoluten Differenzen zwischen der von beiden Systemen ermittelten Anzahl der Spots aufgezeichnet. Beim Vergleich von KS Elispot gegen KS Elispot compact 0.65-Zoom:  $[(R12-R1)]$  Beim Vergleich von KS Elispot gegen KS Elispot compact 1.25-Zoom:  $[(R12-R4)]$ .

In den Bland-Altman-Plots zeigte sich eine Zunahme der Differenzen mit steigender Spotanzahl, so dass die Differenzen folglich nicht als unabhängig von der Spotanzahl angesehen werden konnten und somit auch die Grenzen innerhalb deren 95% der Differenzen zu liegen kommen, nicht parallel zur x-Achse eingezeichnet werden konnten (Vgl. 2. 4. 2). Es wurden daher zusätzlich zu den Bland-Altman-Plots in absoluten Zahlen Bland-Altman-Plots mit prozentualer Auftragung der Differenzen auf der y-Achse gewählt. Durch die Verwendung der relativen Differenzwerte kann die Abhängigkeit der Differenzen von der Spotanzahl aufgehoben werden. Um diese Unabhängigkeit von Differenzen und Spotanzahl zu belegen, wurden die Mittelwerte und die Standardabweichung der Differenzen für einzelne Spotanzahlbereiche berechnet. Es konnte keine Zunahme der Standardabweichung bei steigender Spotanzahl nachgewiesen werden. Die Mittelwerte der Differenzen und die Grenzen innerhalb deren 95% der Differenzen zu erwarten sind, konnten demzufolge parallel zur x-Achse eingezeichnet werden.

Die 95%-Konfidenzgrenzen wurden als relative Werte in Prozent für die gesamte Spotanzahl berechnet. Um jedoch auch Aussagen über kleinere Spotanzahlbereiche treffen zu können, wurden mehrere enge Spotanzahlintervalle gebildet und jeweils aus Mittelwert und Standardabweichung der Differenzen die Grenzen für jeden dieser einzelnen Bereiche in Prozent angegeben. Aus diesen Grenzen konnte nun der prozentuale Anteil der Anzahl der Spots abgelesen werden, die das eine Lesegerät maximal mehr oder weniger als das Vergleichsgerät erkannt hat. Interessierte wie die Geräte bei einer bestimmten Spotanzahl zählten, konnten die 95%-Vertrauensintervalle für den jeweiligen Spotanzahlbereich herangezogen werden.

## ***2.5. Statistische Auswertung***

### **2.5.1. Auswertungszeiten der Bildanalyzesysteme**

Zur Bestimmung der Zeitdauer, die das System KS Elispot und das System KS Elispot compact für die Auswertung einer kompletten Platte benötigt, wurde eine 96-Well-Mikrotiterplatte mit Humanspots an beiden Systemen drei Mal ausgewertet und die dafür am jeweiligen System anfallende Zeit in Minuten gemessen. Es wurde jeweils die reine Lesezeit des automatisierten Systems erfasst.

Während der Messung der Auswertungszeit einer Platte wurde die Zeit für jede der 8 Reihen mit jeweils 12 Wells ermittelt, um so eventuelle Veränderungen der Auswertungsgeschwindigkeit im Verlauf einer Messung zu erfassen.

Um den Einfluss der Spotanzahl auf die Auswertungszeit zu untersuchen, wurde an beiden Systemen bzw. an den drei Messanordnungen die Messzeit für eine Reihe der Platte bestimmt, die Wells mit sehr vielen Spots enthielt und im Gegenzug für eine Reihe, deren Wells nur sehr wenige Spots aufwiesen.

## **2.5.2. Variabilität innerhalb der Messungen**

Um die Variabilität der zwei Systeme zu bestimmen, wurden an jedem Gerät jeweils 50 Wells einer 96-Well-Mikrotiterplatte mit Humanspots zufällig ausgewählt und drei Mal gelesen. Aus den Messergebnissen konnte die Variabilität für beide Systeme errechnet werden. Aus den drei Ergebnissen wurden für jedes Well der Mittelwert, die Standardabweichung und die prozentuale Standardabweichung ermittelt. Aus diesen Werten wurde anschließend die prozentuale Standardabweichung für die Gesamtheit der 50 Wells bestimmt.

## **2.5.3. Zuverlässigkeit der Spoterkennung (Reliability)**

Der Begriff Zuverlässigkeit der Spoterkennung bezieht sich darauf, wie gut das Analysesystem die „wahren“ Spots erkennt und in welchem Ausmaß Farbpunkte, bei denen es sich nicht um „wahre“ Spots handelt, so genannte falsch-positive Spots, in die Messung einbezogen werden.

Dazu wurden für die manuelle Auswertung am System KS Elispot von einzelnen Wells Aufnahmen erstellt und ausgedruckt. Die noch vor der Auswertung durch das System angefertigten Wellaufnahmen wurden unter Verwendung des Bildschirmbildes, das vergrößert werden konnte, sowie des Ausdrucks daraufhin überprüft, bei welchen Farbarealen es sich um „wahre“ Spots handelte und welche keine echten Spots darstellten. Die „wahren“ Spots und die Farbartefakte wurden gekennzeichnet.

Anschließend zählten das System KS Elispot und das System KS Elispot compact mit den zwei Zoom-Einstellungen die Spots in den ausgewählten Wells. Die erkannten und in das Ergebnis eingehenden Spots markierten die Systeme dabei mit einem Ring. Die Wells mit den so als erkannt gekennzeichneten Spots wurden fotografiert und ausgedruckt. Die vom KS Elispot und vom KS Elispot compact ausgewerteten Wellaufnahmen wurden mit den zuvor am KS Elispot erstellten und manuell ausgewerteten Well-Bildern verglichen.

Insgesamt wurden 9 Wells mit Humanspots und einer Spotanzahl bis maximal 200 Spots pro Well zufällig ausgewählt und ausgewertet. Bei einer Spotanzahl unter 200 konnte der Vergleich mit Auge sicher durchgeführt werden. Die Anzahl der an beiden Systemen erkannten „wahren“ Spots und der Anteil der falsch-

positiven Spots am Ergebnis wurden ausgezählt und in absoluten und relativen Zahlen für jedes Well und für die Gesamtheit der Wells angegeben.

#### **2.5.4. Verteilung der Spotgrößen auf Maus-und Humanplatten**

Neben der Anzahl der Spots wurde auch die Größe der Spots erfasst. Die Verteilung der Spotgrößen auf den Platten wurde aus den in der Messanordnung 1 und 2 gewonnenen Daten errechnet. In der spotgrößenabhängigen Messung (Messanordnung 2) wurden für Mausplatten die beiden Größenbereiche [20-100µm], [100-300µm] und für Humanplatten die drei Größenkategorien [20-50µm], [50-100µm], [100-300µm] festgelegt. In die durchmesserunabhängige Messung (Messanordnung 1) wurden alle Spotgrößen einbezogen. Der relative Anteil des jeweiligen Größenbereiches konnte indirekt bestimmt werden:  $\text{Relativer Anteil}(x) = \text{Spotanzahl des Spotgrößenbereiches}(x) / \text{Spotanzahl Gesamtbereich}$ .

Zuvor war zu bestätigen, dass bei Maus-und Humandaten die Summe der Spots der einzelnen Größenbereiche im Bereich der Spotanzahl der Messung lag, die alle Spotgrößen einbezogen hatte. Weiter war zu prüfen, ob eine Abhängigkeit zwischen der Spotgrößenverteilung und der Spotanzahl pro Well bestand. Dazu wurden in einem Diagramm (Scatterplot) die einzelnen relativen Anteile (Größenkategorie 1, 2, 3/ Größenkategorie 4) gegen die Gesamtspotanzahl (Größenkategorie 4) aufgetragen und Regressionslinien eingezeichnet.

#### **2.5.5. Bland-Altman-Plots**

In dieser Arbeit wurden die Ergebnisse des Vergleichs der Systeme KS Elispot und KS Elispot compact mit den zwei Zoom-Einstellungen graphisch mittels Bland-Altman-Plots in absoluten Zahlen und als relative Werte in Prozent mit Angabe der 95%-Konfidenzgrenzen dargestellt. Mittels der Bland-Altman-Plots konnten die Richtigkeit als Mittelwert der Differenzen und die Genauigkeit in Form der 95%-Konfidenzintervalle visualisiert werden (siehe auch 2. 4. 2. und 2. 4. 3). Die Notwendigkeit die Bland-Altman-Plots nicht nur in absoluten sondern auch in relativen Zahlen darzustellen, ergab sich aus der vorliegenden

Unabhängigkeit der Differenzen von der Spotanzahl. In diesem Fall kann eine Auftragung der Konfidenzgrenzen parallel zur Abszisse nur bei prozentualer Skalierung der Ordinate erfolgen.

Die Auswertung der Maus-IFN- $\gamma$ -Platten und der Human-IFN- $\gamma$ -Platten wurde getrennt aufgezeigt. Dabei wurden die Ergebnisse jeder Platte nicht einzeln dargestellt, sondern die Daten der 361 Wells mit Maus-Spots und der 261 Wells mit Humanspots jeweils zusammengefasst. Im Vorfeld wurden alle Elispotplatten einzeln ausgewertet. Dabei konnte kein genereller Unterschied (außer der Spotanzahlverteilung) zwischen den Mausplatten nachgewiesen werden und ebenso war innerhalb der Humanplatten kein Hinweis auf eine Abweichung im Trend zu erkennen.

Zur Darstellung der in der Messanordnung 1 (spotgrößenunabhängige Auswertung) ermittelten Mausdaten wurden vier Diagramme angefertigt:

- (R12-R1) absolut und relativ in % mit Konfidenzgrenzen
- (R12-R4) absolut und relativ in % mit Konfidenzgrenzen

Die in der Messanordnung 2 (spotgrößenabhängige Auswertung) erhobenen Mausdaten wurden in acht Diagrammen zur Darstellung gebracht:

- (R12-R1) 20-100 $\mu\text{m}$  absolut und relativ in % mit Konfidenzgrenzen
- (R12-R1) 100-300 $\mu\text{m}$  absolut und relativ in % mit Konfidenzgrenzen
- (R12-R4) 20-100 $\mu\text{m}$  absolut und relativ in % mit Konfidenzgrenzen
- (R12-R4) 100-300 $\mu\text{m}$  absolut und relativ in % mit Konfidenzgrenzen

Die in der Messanordnung 1 (spotgrößenunabhängige Auswertung) gewonnenen Humandaten wurden in vier Diagrammen dargestellt:

- (R12-R1) absolut und relativ in % mit Konfidenzgrenzen
- (R12-R4) absolut und relativ in % mit Konfidenzgrenzen

Weitere sechs Diagramme dienten zur Visualisierung der in der Messanordnung 2 (spotgrößenabhängige Auswertung) erhobenen Humandaten:

- (R12-R1) 20-50 $\mu\text{m}$ , 50-100 $\mu\text{m}$ , 100-300 $\mu\text{m}$
- (R12-R4) 20-50 $\mu\text{m}$ , 50-100 $\mu\text{m}$ , 100-300 $\mu\text{m}$

Hier wurde aufgrund besonderer Umstände (siehe Ergebnisse) auf die Relativangaben in Prozent verzichtet.

Es wurden die Mittelwerte und die Standardabweichung der Differenzen [(R12-R1); (R12-R4)] für Maus-und Humanspots für die gesamte Spotanzahl und für einzelne Spotanzahlbereiche berechnet und miteinander verglichen. Die Ergebnisse wurden in absoluten Zahlen und als relative Werte in Prozent angegeben.

Die 95%-Vertrauensintervalle wurden für den Vergleich von R12 gegen R1 sowie von R12 gegen R4 bei Maus-und Humanspots sowohl für die gesamte Spotanzahl als auch für einzelne Spotanzahlbereiche angegeben.

## **3. Ergebnisse**

### ***3.1. Auswertungszeiten der Bildanalysesysteme***

Am System KS Elispot wurden 96 Wells einer Mikrotiterplatte mit Humanspots drei Mal gelesen. Der Mittelwert der drei gemessenen Auswertungszeiten betrug 32 Minuten 45 Sekunden (STD 0,05). Die kürzeste Messzeit lag bei 32 Minuten 30 Sekunden, die längste betrug 32 Minuten 50 Sekunden. Für die Messdauer einer der 8 Reihen der 96-Well-Mikrotiterplatte, die jeweils 12 Wells enthält, ergab sich daraus ein durchschnittlicher Wert von 4 Minuten 9 Sekunden.

12 Wells in einer Reihe, die Spotanzahlen von 0 bis 100 pro Well aufwiesen (Mittelwert 58 Spots pro Well), wurden im Mittel in 3 Minuten 53 Sekunden gemessen (STD 0,03). Die kürzeste Messzeit betrug 3 Minuten 50 Sekunden, die längste Messzeit 3 Minuten 56 Sekunden. Der Mittelwert der drei Messzeiten für eine Reihe von 12 Wells, die pro Well 150 bis 900 Spots enthielten (Mittelwert 460 Spots pro Well) war 4 Minuten 48 Sekunden (STD 0,09). Die kürzeste Messzeit betrug 4 Minuten 38 Sekunden, die längste dauerte 4 Minuten 56 Sekunden (Tab.5).

Das System KS Elispot compact 0.65-Zoom (R1) wertete dieselben 96 Wells einer Mikrotiterplatte in 5 Minuten 1 Sekunde aus (STD 0,06). Die kürzeste der drei Messzeiten betrug 5 Minuten, die längste 5 Minuten 12 Sekunden. Die Auswertungszeit einer Reihe aus 12 Wells dauerte demnach durchschnittlich 38 Sekunden.

Eine Reihe aus 12 Wells, die pro Well 0 bis 100 Spots enthielten, wertete das System in 36 Sekunden aus (STD 0,05). Die kürzeste Messzeit war 36 Sekunden, die längste lag bei 37 Sekunden. Der Mittelwert der drei Messzeiten für 12 Wells mit Spotzahlen von 150 bis 900 pro Well betrug 52 Sekunden (STD 0,02). Die kürzeste Messzeit dauerte 50 Sekunden, die längste ergab 54 Sekunden (Tab. 5).

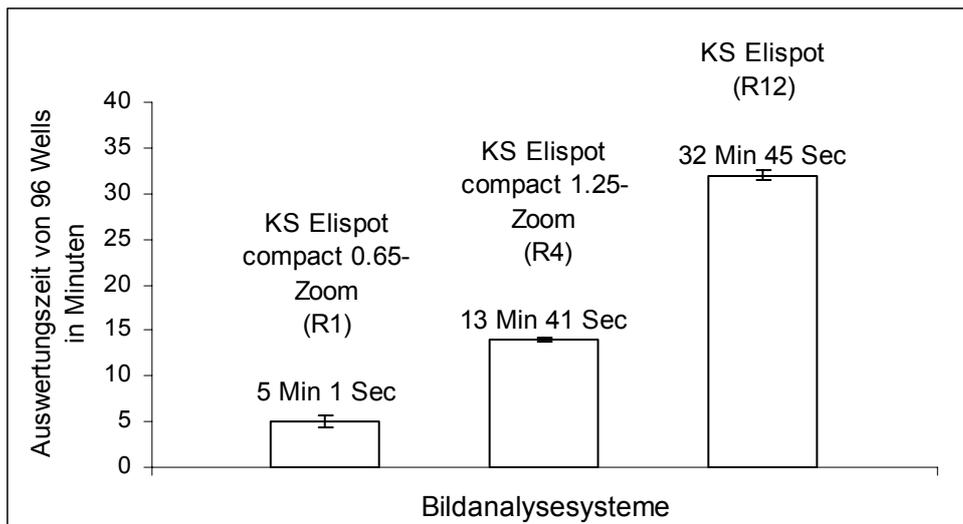
Am System KS Elispot compact 1.25-Zoom (R4) ergab der Mittelwert der drei Auswertungszeiten der 96 Wells einer Mikrotiterplatte 13 Minuten 41 Sekunden (STD 0.03). Die kürzeste Zeit betrug 13 Minuten 38 Sekunden, die längste Zeit

lag bei 13 Minuten 44 Sekunden. Für eine aus 12 Wells bestehende Reihe ergab sich eine durchschnittliche Auswertungszeit von 1 Minute 42 Sekunden.

12 Wells in einer Reihe, die Spotanzahlen von 0 bis 100 pro Well aufwiesen, wurden in 1 Minute 32 Sekunden ausgewertet (STD 0.02). Die kürzeste Zeit war 1 Minute 30 Sekunden, die längste Auswertungszeit betrug 1 Minute 35 Sekunden. Die Auswertungszeit einer Reihe von 12 Wells, die pro Well 150 bis 900 Spots enthielten, war im Mittel 2 Minuten 6 Sekunden (STD 0.04). Die kürzeste Messzeit lag bei 2 Minuten 2 Sekunden, die längste dauerte 2 Minuten 11 Sekunden (Tab.5).

| Bildanalysesystem                        | Auswertungszeit von 96 Wells | Auswertungszeit einer 12-Well-Reihe mit 0-100 Spots/ Well | Auswertungszeit einer 12-Well-Reihe mit 150-800 Spots/ Well |
|--|------------------------------|---|---|
| <b>KS Elispot (R12)</b>                  | 32 Min 45 Sec                | 3 Min 53 Sec  | 4 Min 48 Sec  |
| <b>KS Elispot compact 0.65-Zoom (R1)</b> | 5 Min 1 Sec                  | 36 Sec  | 52 Sec  |
| <b>KS Elispot compact 1.25-Zoom (R4)</b> | 13 Min 41 Sec                | 1 Min 32 Sec  | 2 Min 6 Sec   |

**Tab. 5:** Auswertungszeiten der verglichenen Bildanalysesysteme. Bei Verwendung eines Bildes pro Well (0.65-Zoom) wertete das KS Elispot compact eine komplette 96-Well-Mikrotiterplatte bis zu 6 Mal schneller aus als das KS Elispot, bei Verwendung von 4 Bildern pro Well (1.25-Zoom) war das KS Elispot compact im Vergleich zum KS Elispot mindestens doppelt so schnell.



**Abb. 17:** Vergleich der Auswertungszeiten von 96 Wells am System KS Elispot und KS Elispot compact (0.65-Zoom und 1.25-Zoom). Das System KS Elispot wies die längste Auswertungszeit für eine komplette Mikrotiterplatte auf, das System KS Elispot compact 0.65-Zoom wertete eine Mikrotiterplatte in der kürzesten Zeit aus.

### 3.2. Variabilität innerhalb der Messungen

Aus einer 96-Well-Mikrotiterplatte mit Humanspots wurden 50 Wells mit einer Spotanzahl von 10 bis 900 Spots pro Well zufällig ausgewählt und drei Mal an beiden Systemen bzw. an den drei Messanordnungen gemessen. Aus den drei Werten pro Well wurden der Mittelwert, die Standardabweichung und die Variabilität in % für jedes Well und für die 2 Systeme bzw. für die 3 Messanordnungen berechnet. Aus den Einzelergebnissen wurde die Variabilität für die Gesamtheit der Wells berechnet und für die einzelnen Systeme angegeben (Tab. 6).

| <b>Bildanalysesystem</b>             | <b>Variabilität innerhalb der Messungen</b> |
|--------------------------------------|---|
| KS Elispot<br>(R12)                  | 2.30%                                       |
| KS Elispot compact 0.65-Zoom<br>(R1) | 2.68%                                       |
| KS Elispot compact 1.25-Zoom<br>(R4) | 2.73%                                       |

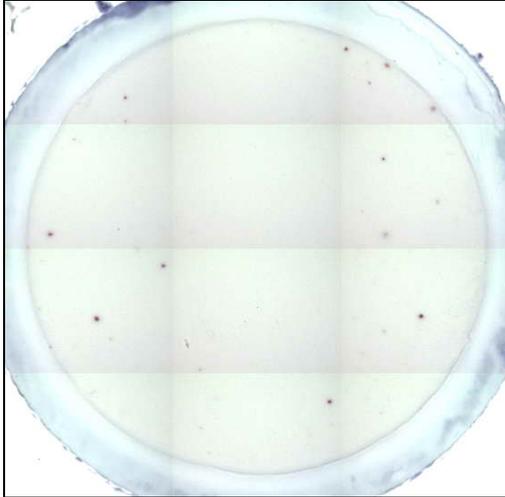
**Tab. 6:** Variabilität innerhalb der Messungen. Die verglichenen Systeme wiesen niedrige Variabilitätswerte ohne nennenswerte Unterschiede auf.

### **3.3. Zuverlässigkeit der Spoterkennung (Reliability)**

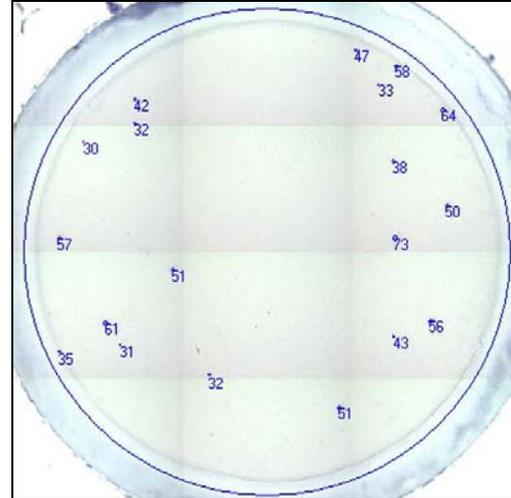
Unter dem Begriff Zuverlässigkeit der Spoterkennung war zu prüfen, wie gut die Readersysteme die „wahren“ Spots erkannten und in welchem Ausmaß durch Schmutz oder Artefakte entstandene Farbareale, bei denen es sich nicht um „wahre“ Spots handelte, als so genannte falsch-positive Spots in die Messung einbezogen wurden.

Dazu wurden 9 Wells mit einer Spotanzahl bis maximal 200 Spots pro Well zufällig ausgewählt. Bei einer Spotanzahl unter 200 pro Well war der Vergleich mit dem Auge sicher durchführbar. Zunächst wurden am System KS Elispot von den 9 ausgewählten Wells Aufnahmen angefertigt und ausgedruckt (Abb. 18a). Unter Verwendung des Ausdruckes und des Bildschirmbildes, das vergrößert werden konnte, wurde mit dem Auge überprüft, bei welchen Farbpunkten es sich um „wahre“ Spots handelte und welche keine echten Spots darstellten. Diese Aufnahmen wurden anschließend mit den, vom KS Elispot und vom KS Elispot compact ausgewerteten und ausgedruckten Wellaufnahmen verglichen. Auf den nach Auswertung angefertigten Aufnahmen waren die vom System erkannten Spots mit einem Farbring markiert (Abb. 18b, 19b, c). Es wurde die Anzahl der erkannten echten Spots, die Anzahl der nicht erkannten Spots, sowie die Anzahl der falsch-positiven Spots am Ergebnis für das System KS Elispot und für das

System KS Elispot compact mit den zwei Auflösungen (0.65-Zoom und 1.25-Zoom) bestimmt.



**Abb. 18a**



**Abb. 18b**

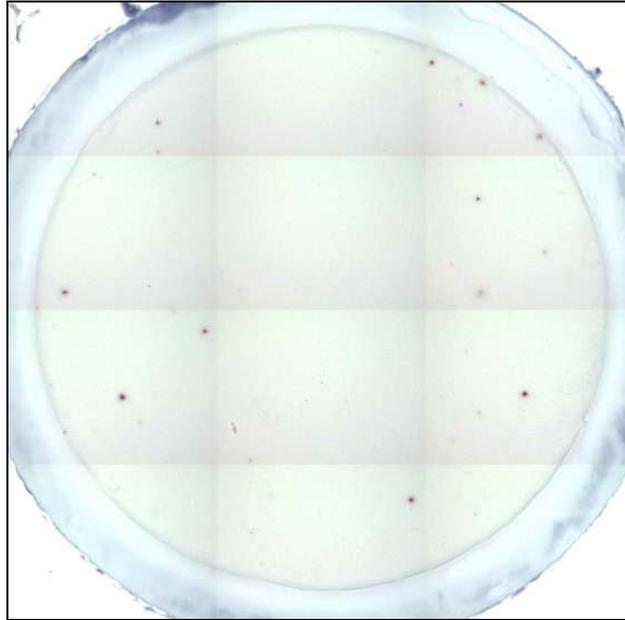
**Abb. 18a, b:** Beispielhafte Aufnahmen eines Wells (Well 3) mit Humanspots. Abbildung a wurde vor Auswertung, Abbildung b nach Auswertung des Wells durch das System aufgenommen. Die vom KS Elispot erkannten Spots wurden mit einem Farbring markiert, die daneben stehenden Zahlen geben den Spotdurchmesser an. Das System zählte 19 Spots, manuell wurden 20 Spots erfasst.

Das System KS Elispot (R12) hatte insgesamt 95.5% der echten Spots identifiziert. 0% waren falsch-positiv. 4.5% der „wahren“ Spots wurden von diesem Reader-System nicht erkannt (Tab. 7).

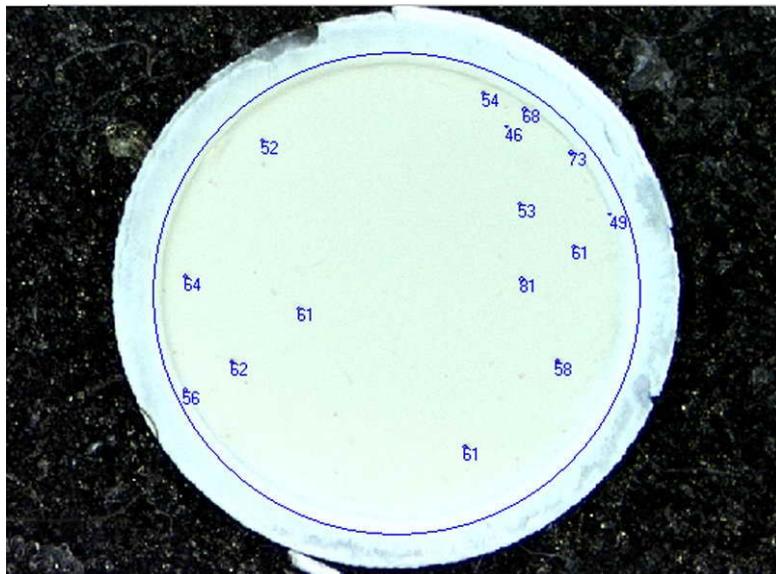
| <b>KS<br/>Elispot<br/>(R12)</b> | Anzahl<br>wahrer<br>Spots | Ergebnis<br>KS<br>Elispot<br>(R12) | Anzahl<br>der<br>erkannt.<br>wahren<br>Spots | Anzahl<br>der nicht<br>erkannt.<br>wahren<br>Spots | Anteil<br>falsch-<br>positiver<br>Spots am<br>Ergebnis |
|---------------------------------|---------------------------|------------------------------------|--|--|--|
| Well 1                          | 2                         | 2                                  | 2<br>(100%)                                  | 0<br>(0%)  | 0<br>(0%)  |
| Well 2                          | 10                        | 10                                 | 10<br>(100%)                                 | 0<br>(0%)  | 0<br>(0%)  |
| Well 3                          | 20                        | 19                                 | 19<br>(95%)                                  | 1<br>(5%)  | 0<br>(0%)  |
| Well 4                          | 27                        | 25                                 | 25<br>(92,59%)                               | 2<br>(7,41%)                                       | 0<br>(0%)  |
| Well 5                          | 51                        | 47                                 | 47<br>(92,16%)                               | 4<br>(7,84%)                                       | 0<br>(0%)  |
| Well 6                          | 58                        | 55                                 | 55<br>94,83%)                                | 3<br>(5,17%)                                       | 0<br>(0%)  |
| Well 7                          | 93                        | 89                                 | 89<br>(95,7%)                                | 4<br>(4,3%)  | 0<br>(0%)  |
| Well 8                          | 109                       | 104                                | 104<br>(95,41%)                              | 5<br>(4,59%)                                       | 0<br>(0%)  |
| Well 9                          | 162                       | 152                                | 152<br>(93,83%)                              | 10<br>(6,17%)                                      | 0<br>(0%)  |
| Gesamt                          |                           |                                    | 95,5%  | 4,5%   | 0%   |

**Tab. 7:** Anzahl der vom KS Elispot (R12) erkannten und nicht erkannten „wahren“ Spots, sowie Anteil der falsch-positiven Spot am Ergebnis (absolut bzw. in %). Das System KS Elispot erkannte bemerkenswerterweise keine falsch-positiven Spots.

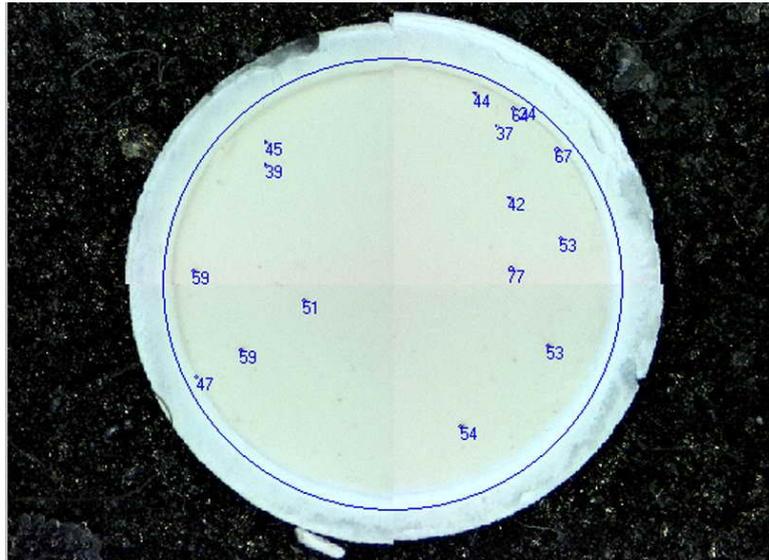
Für die Ermittlung der Zuverlässigkeit des Systems KS Elispot compact (R1 und R4) wurden die 9 Wellbilder nach Auswertung durch das System aufgenommen, ausgedruckt und mit den am System KS Elispot erstellten Aufnahmen manuell verglichen (Abb. 19a-c). Analog wurden die Anzahl der erkannten echten Spots, die Anzahl der nicht erkannten Spots, sowie der Anteil der falsch-positiven Spots an den Ergebnissen von R1 und R4 berechnet (Tab. 8, 9).



**Abb. 19a**



**Abb. 19b**



**Abb. 19c**

**Abb. 19a-c:** Beispielhafte Aufnahmen desselben Wells (Well 3) mit Humanspots, erstellt am KS Elispot (Abb. 19a) und am KS Elispot compact 0.65-Zoom (Abb. 19b) bzw. 1.25-Zoom (Abb. 19c). Die vom System KS Elispot compact erkannten Spots wurden mit einem Farbring markiert und mit der am KS Elispot erstellten Well-Aufnahme verglichen. Die Zahlen neben der Markierung geben den Durchmesser an. Das System KS Elispot compact 0.65-Zoom zählte 14 „wahre“ Spots (19b), das System KS Elispot compact 1.25-Zoom maß 15. Manuell wurden in diesem Well 20 echte Spots erfasst (19a).

Das KS Elispot compact 0.65-Zoom (R1) erkannte 79.36% der echten Spots korrekt. 20.64% der echten Spots wurden von diesem System nicht erkannt. Der Anteil der falsch-positiven Spots an den Ergebnissen von R1 betrug 5.85% (Tab. 8).

| <b>KS Elispot compact (R1)</b> | Anzahl der wahren Spots | Ergebnis KS Elispot compact (R1) | Anzahl der erkannt. wahren Spots | Anzahl der nicht erkannt. wahren Spots | Anteil falsch-positiver Spots am Ergebnis |
|--------------------------------|-------------------------|----------------------------------|----------------------------------|--|---|
| Well 1                         | 2                       | 2                                | 2<br>(100%)                      | 0<br>(0%)                              | 0<br>(0%)                                 |
| Well 2                         | 10                      | 8                                | 7<br>(70%)                       | 3<br>(30%)                             | 1<br>(12,5%)                              |
| Well 3                         | 20                      | 15                               | 14<br>(70%)                      | 6<br>(30%)                             | 1<br>(6,7%)                               |
| Well 4                         | 27                      | 20                               | 18<br>(66.7%)                    | 9<br>(33.3%)                           | 2<br>(10%)                                |
| Well 5                         | 51                      | 42                               | 39<br>(76.47%)                   | 12<br>(23.53%)                         | 3<br>(7.14%)                              |
| Well 6                         | 58                      | 49                               | 46<br>(79.31%)                   | 12<br>(20.69%)                         | 3<br>(6.12%)                              |
| Well 7                         | 93                      | 81                               | 77<br>(82.79%)                   | 16<br>(17.2%)                          | 4<br>(4.9%)                               |
| Well 8                         | 109                     | 95                               | 92<br>(84.4%)                    | 17<br>(15.6%)                          | 3<br>(3.16%)                              |
| Well 9                         | 162                     | 140                              | 137<br>(84.57%)                  | 25<br>(15.43%)                         | 3<br>(2.14%)                              |
| Gesamt                         |                         |                                  | 79.36%                           | 20.64%                                 | 5.85%                                     |

**Tab. 8:** Anzahl der vom KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) erkannten und nicht erkannten „wahren“ Spots, sowie der Anteil der falsch-positiven Spots am Gesamtergebnis R1 (absolut bzw. in %). Auffällig ist der bei diesem System höchste Anteil an falsch-positiven Spots.

Die Auswertung der am KS Elispot compact 1.25-Zoom (R4) aufgenommenen Well-Bilder ergab, dass dieses System 84.29% der „wahren“ Spots richtig erfasst hatte. 15.71% der „wahren“ Spots wurden nicht erkannt. Der Anteil der falsch-positiven Spots an den Ergebnissen von R4 betrug 3.69% (Tab. 9).

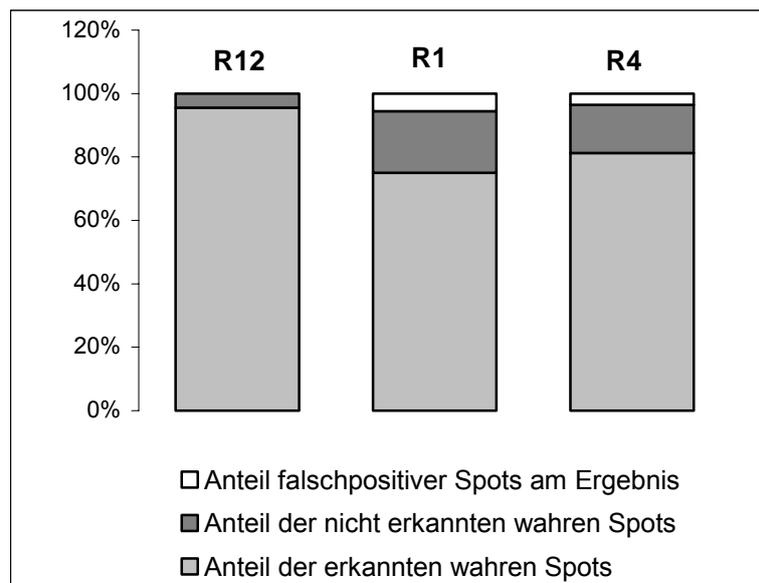
| KS Elispot compact (R4) | Anzahl der Wahren Spots | Ergebnis KS Elispot compact (R4) | Anzahl der erkannt. wahren Spots | Anzahl der nicht erkannt. wahren Spots | Anteil falsch-positiver Spots am Ergebnis |
|-------------------------|-------------------------|----------------------------------|----------------------------------|--|---|
| Well 1                  | 2                       | 2                                | 2<br>(100%)                      | 0<br>(0%)                              | 0<br>(0%)                                 |
| Well 2                  | 10                      | 8                                | 8<br>(80%)                       | 2<br>(20%)                             | 0<br>(0%)                                 |
| Well 3                  | 20                      | 16                               | 15<br>(75%)                      | 5<br>(25%)                             | 1<br>(6,25%)                              |
| Well 4                  | 27                      | 21                               | 20<br>(74.07%)                   | 7<br>(25.93%)                          | 1<br>(4,76%)                              |
| Well 5                  | 51                      | 44                               | 40<br>(78.43%)                   | 11<br>(21.57%)                         | 4<br>(9,09%)                              |
| Well 6                  | 58                      | 52                               | 50<br>(86.21%)                   | 8<br>(13.79%)                          | 2<br>(3.85%)                              |
| Well 7                  | 93                      | 85                               | 82<br>(88.17%)                   | 11<br>(11.83%)                         | 3<br>(3.53%)                              |
| Well 8                  | 109                     | 99                               | 94<br>(86.24%)                   | 15<br>(13.76%)                         | 5<br>(5.05%)                              |
| Well 9                  | 162                     | 147                              | 146<br>(90.12%)                  | 16<br>(9.88%)                          | 1<br>(0.6%)                               |
| Gesamt                  |                         |                                  | 84.29%                           | 15.71%                                 | 3.69%                                     |

**Tab. 9:** Anzahl der vom KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) erkannten und nicht erkannten „wahren“ Spots, sowie Anteil der falsch-positiven Spots am Gesamtergebnis R4 (absolut bzw. in %).

Die zur Beurteilung der Zuverlässigkeit der Spoterkennung für beide Reader-Systeme bzw. für die 3 Messanordnungen ermittelten Ergebnisse wurden in Tabelle 10 zusammengefasst. Der Anteil der erkannten echten Spots lag beim KS Elispot höher als beim KS Elispot compact 0.65-Zoom (R1) und 1.25-Zoom (R4). Am niedrigsten war die Anzahl der erkannten echten Spots beim KS Elispot compact 0.65-Zoom (R1). Das KS Elispot wies als einziges System keine falsch-positiven Spots am Ergebnis auf, während das KS Elispot compact 0.65-Zoom (R1) den höchsten Anteil an falsch-positiven Spots zeigte (Abb. 20).

|   | KS Elispot (R12) | KS Elispot compact 0.65-Zoom (R1) | KS Elispot compact 1.25-Zoom (R4) |
|---|------------------|-----------------------------------|-----------------------------------|
| Anzahl der erkannten wahren Spots in 9 Wells in %       | 95.5%            | 79.36%                            | 84.29%                            |
| Anzahl der nicht erkannten wahren Spots in 9 Wells in % | 4.5%             | 20.64%                            | 15.71%                            |
| Anzahl der falsch-positiven Spots in 9 Wells in %       | 0%               | 5.85%                             | 3.69%                             |

**Tab. 10:** Vergleich der Ergebnisse der Bildanalyseysteme KS Elispot, KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) und KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bezüglich der Zuverlässigkeit der Spoterkennung.



**Abb. 20:** Der Anteil der erkannten „wahren“ Spots lag beim KS Elispot (R12) mit 95.5% am höchsten, beim KS Elispot compact 0.65-Zoom (R1) war er mit 79.36% am niedrigsten, das KS Elispot 1.25-Zoom (R4) erkannte 84.29% der „wahren“ Spots. Das KS Elispot wies im Gegensatz zum KS Elispot compact 0.65-Zoom (5.85%) und zum KS Elispot compact 1.25-Zoom (3.69%) keine falsch-positiven Spots am Ergebnis auf.

### **3.4. Spotgrößenverteilung**

Die Spotanzahl pro Well erstreckte sich bei den Maus-IFN- $\gamma$ -Spots von 0 bis 283. 70% der Mausspots lagen im Größenbereich 20-100 $\mu$ m, 30% fielen in den Größenbereich 100-300 $\mu$ m.

Bei den Human-IFN- $\gamma$ -Spots enthielten die Wells 0 bis 906 Spots pro Well. Von den Humanspots verteilten sich 76.3% auf den Bereich 20-50 $\mu$ m, 21.4% auf den Bereich 50-100 $\mu$ m und 2.28% auf den Bereich 100-300 $\mu$ m.

Die einzelnen relativen Größenanteile der Human-und Mausspots wurden in einem Scatterplot gegen die Gesamtspotanzahl aufgetragen. Nach Einzeichnung einer Regressionslinie in die einzelnen Diagramme, ergab sich in allen Fällen eine Gerade, die fast parallel zur x-Achse, der Spotanzahlachse, verlief. Das Resultat berechtigte zur Annahme, dass die relativen Anteile der drei Spotgrößenkategorien bei Humanspots, sowie die relativen Anteile der zwei Spotgrößeneinteilungen bei Mausspots, unabhängig von der Spotanzahl waren.

### **3.5. Ergebnisse der Auswertung von Mausspots**

#### **3.5.1. Mittelwerte der Differenzen für Mausspots**

Die Mittelwerte (MW) aller Differenzen (R12-R1) und (R12-R4) wurden für die gesamte Spotanzahl sowie für einzelne, kleinere Spotanzahlintervalle berechnet. Sie wurden in absoluten Zahlen und als relative Werte in Prozent mit Standardabweichung (STD) angegeben. Eine Differenz (R12-R1) bzw. (R12-R4), die größer als null war bedeutete, dass die von R12 gemessene Spotanzahl über der von R1 bzw. R4 lag.

| Spotanzahl n<br>[Mw(R12;R1);<br>Mw(R12;R4)] | MW<br>(R12-R1)<br>absolut | STD             | MW<br>(R12-R4)<br>absolut | STD   |
|---|---------------------------|-----------------|---------------------------|-------|
| n gesamt                                    | 3.62                      | 18.22           | 1.54                      | 6.76  |
| [0-100]                                     | -0.17                     | 5.28            | 1.14                      | 3.27  |
| [0-50]                                      | -1.10                     | 3.04            | 0.87                      | 2.10  |
| [50-100]                                    | 12.31                     | 10.87           | 4.47                      | 8.92  |
| [100-200]                                   | 44.57<br>33.57*           | 41.42<br>19.08* | 3.58                      | 18.00 |

**Tab. 11:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) in absoluten Zahlen bei Mausspots.

\*ohne zwei Ausreißerwerte

| Spotanzahl n<br>[Mw(R12;R1);<br>Mw(R12;R4)] | MW<br>(R12-R1)<br>relativ | STD               | MW<br>(R12-R4)<br>relativ | STD    |
|---|---------------------------|-------------------|---------------------------|--------|
| n <sub>≥</sub> 3                            | -4.14%                    | 0.3546            | 10.29%                    | 0.2176 |
| [3-100]                                     | -10.61%                   | 0.3314            | 11.88%                    | 0.2328 |
| [3-50]                                      | -15.13%                   | 0.3340            | 12.82%                    | 0.2453 |
| [50-100]                                    | 16.23%                    | 0.1277            | 6.14%                     | 0.1212 |
| [100-200]                                   | 30.76%<br>24.14%*         | 0.2585<br>0.1373* | 2.65%                     | 0.1152 |

**Tab. 12:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) in relativen Zahlen bei Mausspots.

\* ohne zwei Ausreißerwerte

Der Mittelwert (MW) aller Differenzen (R12-R1) [absolut bzw. in %] betrug [3.62 bzw. -4.14%], der Mittelwert aller Differenzen (R12-R4) hatte den Wert [1.54 bzw. 10.29%] (Tab. 11, 12). Bei der Bestimmung der Differenzmittelwerte in relativen Zahlen wurden Wells mit einer Spotanzahl von drei oder mehr Spots pro Well in die Auswertung einbezogen, da die prozentuale Auftragung bei kleineren Spotzahlen zu großen Ausschlägen und damit zur Verzerrung des Gesamtergebnisses geführt hätte.

Die Mittelwerte der Differenzen (R12-R1) und (R12-R4) wurden auch für einzelne, engere Spotanzahlintervalle angegeben, um so Unterschiede der Messwerte abhängig von der Spotanzahl zu erkennen.

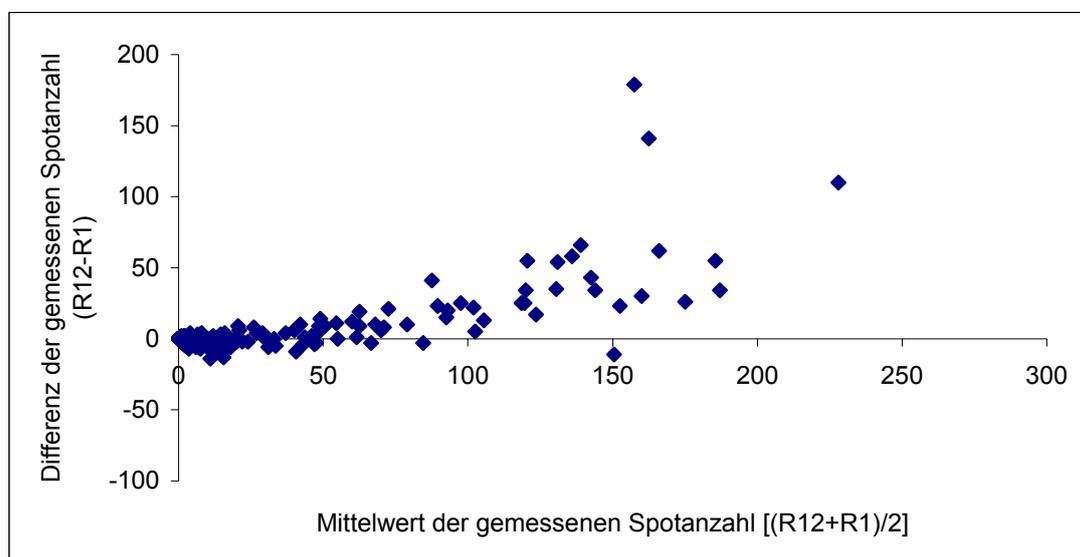
Im Intervall [0-100] war der Mittelwert aller Differenzen (R12-R1) mit [-0.17 bzw. -10.61%] kleiner als der Mittelwert aller Differenzen (R12-R4) mit [1.14 bzw. 11.88%] und wies negative Werte auf. Der negative Differenzmittelwert (R12-R1) bedeutete, dass R12 in diesem Intervall überwiegend weniger Spots als R1 gemessen hatte, der positive Differenzmittelwert (R12-R4) sagte aus, dass R12 mehrheitlich mehr Spots als R4 erkannt hatte. Die weitere Unterteilung dieses Intervalls in die Bereiche [0-50] und [50-100] Spots pro Well zeigte, dass der negative Mittelwert der Differenzen (R12-R1) im Intervall [0-100] hauptsächlich auf den Differenzmittelwert (R12-R1) im Spotanzahlbereich [0-50] zurückzuführen war, dieser betrug [-1.10 bzw. -15.13%], der Differenzmittelwert (R12-R4) war [0.87 bzw. 12.82%]. Im nächst höheren Intervall [50-100] wiesen sowohl der Differenzmittelwert (R12-R1) [12.31 bzw. 16.23%] als auch der Differenzmittelwert (R12-R4) [4.47 bzw. 6.14%] positive Werte auf, R12 zählte folglich überwiegend mehr Spots als R1 und als R4. Der Mittelwert der Differenzen (R12-R1) war dabei mehr als doppelt so groß wie der Mittelwert der Differenzen (R12-R4) (Tab. 11, 12). R1 maß demnach weniger Spots als R4.

Die größten Unterschiede zwischen den Mittelwerten der Differenzen (R12-R1) und (R12-R4) wurden bei Spotanzahlen über 100 pro Well gesehen. Im Intervall [100-200] lag der Mittelwert aller Differenzen (R12-R1) mit [33.57 bzw. 24.14%] klar über dem Mittelwert aller Differenzen (R12-R4), der [3.58 bzw. 2.65%] betrug. R12 zählte folglich überwiegend mehr Spots als R1 und als R4. R1 erkannte weniger Spots als R4.

### 3.5.2. Bland-Altman-Plots für Mausspots in absoluten Zahlen

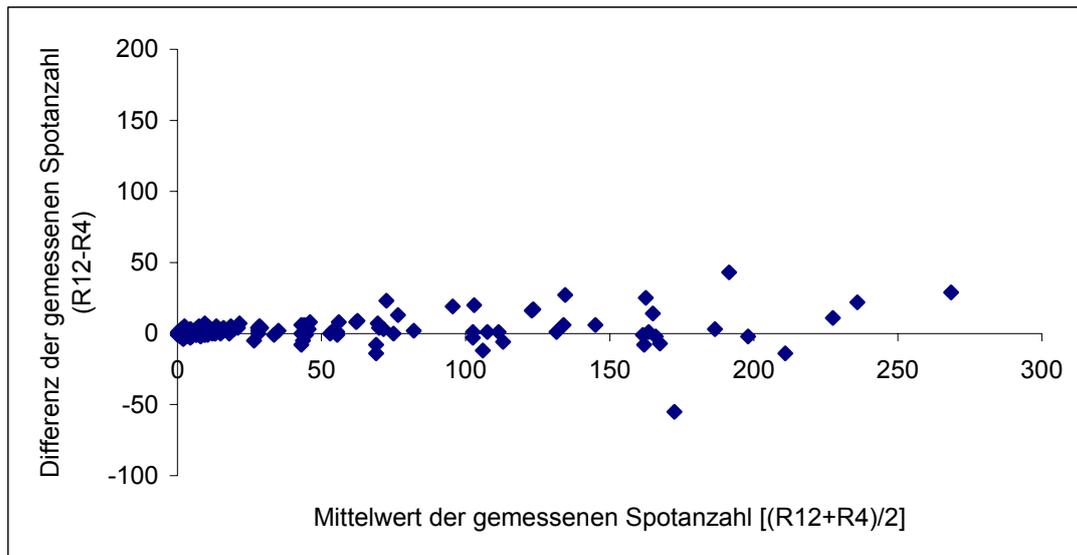
In den Bland-Altman-Plots wurden auf der Abszisse die Mittelwerte der von beiden Systemen pro Well gezählten Spotanzahl aufgetragen. Auf der Ordinate wurden die absoluten Differenzen zwischen der von beiden Systemen ermittelten Anzahl der Spots aufgezeichnet (Abb. 21, 22). Jeder Punkt im Diagramm repräsentierte somit ein durch beide Systeme ausgewertetes Well. Beim Vergleich R12 gegen R1 bzw. R12 gegen R4 bedeutete demnach ein oberhalb der x-Achse liegender Wert, dass R12 in diesem Well mehr Spots als R1 bzw. R4 gezählt hatte. Entsprechend drückte ein unterhalb der x-Achse gelegener Punkt aus, dass R12 in diesem Well weniger Spots als R1 bzw. R4 gezählt hatte. Je größer die Werte waren, desto größer war der Unterschied zwischen der von R12 und R1 bzw. R4 gemessenen Spotanzahl. Ein exakt auf der Abszisse liegender Wert bedeutete, dass die von beiden Systemen ermittelten Ergebnisse übereinstimmten.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Mausspots**



**Abb. 21:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12, 12 Bilder pro Well) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Mausspots. Die meisten Differenzen ( $R12-R1$ ), im Diagramm als Rauten dargestellt, waren positiv. R12 zählte demnach überwiegend mehr Spots als R1.

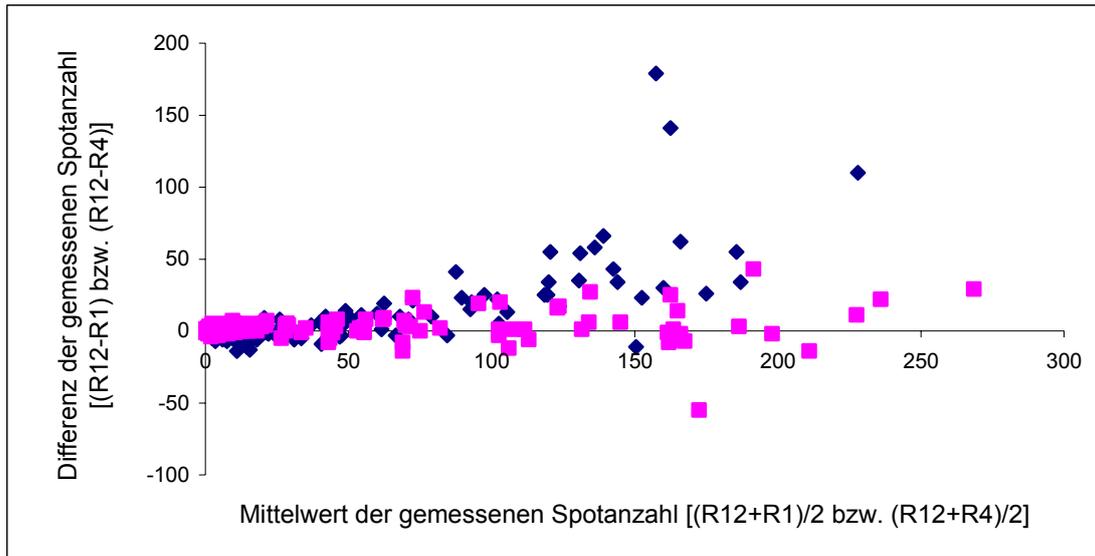
## KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei Mausspots



**Abb. 22:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12, 12 Bilder pro Well) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Mausspots. Die Mehrzahl der Differenzen (R12-R4), im Diagramm als Rauten dargestellt, hatte positive Werte. R12 zählte demnach überwiegend mehr Spots als R4.

Bei der Beurteilung der Bland-Altman-Plots, die den Vergleich von R12 mit R1 und R12 mit R4 graphisch abbildeten, waren der Spotanzahlbereich [0-100] und die darüber liegenden Bereiche zu unterscheiden.

**KS Elispot (R12) im Vergleich zu KS Elispot compact (R1, R4) bei  
Mausspots**



**Abb. 23:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1) (veranschaulicht als Rauten  $\diamond$ ), sowie für KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4) (veranschaulicht als Quadrate  $\square$ ) bei Mausspots. Auffällig sind die Unterschiede zwischen R1 und R4 bei Spotzahlen über 100 pro Well.

Die Bland-Altman-Plots zeigten für den Vergleich R12 mit R1 und R12 mit R4 bei einer Spotanzahl unter 100 weitgehend einen ähnlichen graphischen Verlauf, während sie bei Spotanzahlen über 100 pro Well stärker differierten (Abb. 23).

Im Intervall [0-100] Spots pro Well waren Unterschiede zum einen im Spotanzahlbereich [0-50] nachzuweisen: in diesem Abschnitt divergierten (R12-R1) und (R12-R4) in ihren Anteilen an negativen Differenzen. Eine negative Differenz bedeutete, dass R12 weniger Spots als R1 bzw. als R4 zählte. (R12-R1) wies eine ca. dreifach höhere Anzahl an negativen Differenzwerten als (R12-R4) auf: bei (R12-R1) lagen im Spotbereich [0-50] insgesamt 257 Differenzen im Intervall [-14; 14]. 161 der Differenzen und damit 62.3% waren kleiner als 0. 96 Differenzen lagen im positiven Wertebereich. Demnach maß R12 in diesem Bereich überwiegend weniger Spots als R1. Im Gegensatz zu den Differenzen (R12-R1) waren die meisten der 207 Differenzwerte (R12-R4) positiv. 164 der

Differenzen waren größer oder gleich null, 43 und damit 20.7% hatten einen negativen Wert. Damit zählte R12 in diesem Bereich meist mehr Spots als R4. Die Differenzen (R12-R4) verteilten sich im Intervall [-8; 8].

Zum anderen waren im Spotanzahlbereich [50-100] Unterschiede zwischen (R12-R1) und (R12-R4) festzustellen: die 19 Differenzen (R12-R1), die sich im Intervall [-3; 41] befanden, waren in 17 Fällen positiv, zwei von ihnen hatten negative Werte. Damit zählte R12 fast immer mehr Spots als R1. Bei (R12-R4) zählte man im entsprechenden Bereich 17 Differenzen im Intervall [-14; 23]. 14 der Differenzwerte lagen oberhalb der x-Achse, 3 waren negativ und dementsprechend unterhalb der x-Achse zu finden. R12 zählte somit fast immer mehr Spots als R4. Die positiven Differenzwerte bei (R12-R1) waren höher als die Werte bei (R12-R4). Damit zählte R1 weniger Spots pro Well als R4. Bei (R12-R4) betrug die Maximaldifferenz im Positivbereich 23. 11 der 14 positiven Differenzwerte (R12-R4), und damit 79%, waren kleiner als 10. Die Differenzen (R12-R1) erreichten im Positivbereich den Maximalwert 41. Dieser betrug damit fast das Doppelte des Maximalwertes (R12-R4). 4 der 17 positiven Differenzen (R12-R1), damit 23.5%, wiesen Werte unter 10 auf, die restlichen kamen darüber zu liegen.

Bei Spotanzahlen über 100 pro Well wichen (R12-R1) und (R12-R4) deutlicher in der graphischen Darstellung voneinander ab (Abb. 23).

Bei Spotzahlen über 100 pro Well konnten 23 Differenzen (R12-R1) im Intervall [-11; 179] nachgewiesen werden. Einer der 23 Werte war negativ. Von den 24 Differenzen (R12-R4) waren 9 negativ, sie lagen im Intervall [-55; 43]. Die Anzahl der positiven Differenzwerte überwog, R12 zählte demnach überwiegend mehr Spots als R1 und als R4. Die Differenzwerte (R12-R1) erreichten im Spotanzahlbereich [100-200] höhere Werte im positiven Zahlenbereich als die Differenzen (R12-R4). R1 maß damit weniger Spots pro Well als R4. Während bei einer Spotanzahl unter 100 die Differenzen (R12-R1) maximal den Wert 41 zeigten, stiegen die Differenzwerte (R12-R1) im Bereich [100-200] Spots pro Well auf Maximalwerte von bis zu 179 an. Insgesamt hatte die Hälfte der Differenzen (R12-R1) Werte von 50 bis 179 (unter Vernachlässigung der zwei Ausreißerwerte 141 und 179, Werte von 50 bis 70). Die übrigen lagen bis auf den Wert [-11] im Wertebereich [0-50]. Bei (R12-R4) waren hingegen alle

Differenzwerte kleiner als 50 und zeigten den bei (R12-R1) beschriebenen ausgeprägten Anstieg der Werte im Positivbereich gegenüber dem darunter liegenden Spotanzahlbereich nicht.

### **3.5.3. Bland-Altman-Plots für Mausspots in relativen Zahlen**

Die Messdaten aus dem Vergleich R12 gegen R1 und R12 gegen R4 wurden mittels Bland-Altman-Plots auch in relativen Zahlen graphisch veranschaulicht (Abb. 24, 25). Auf der Abszisse wurden analog zum Vorgehen bei den absoluten Zahlen die Mittelwerte der von beiden Systemen pro Well gezählten Spotanzahl aufgetragen. Auf der Ordinate wurden die Differenzen zwischen der von beiden Systemen ermittelten Anzahl der Spots in Prozent aufgetragen. Sind, wie im vorliegenden Fall, die Differenzen nicht unabhängig von den Werten auf der x-Achse, so dass die Mittelwerte und die Standardabweichung der Differenzen mit zunehmenden x-Werten ansteigen, können die 95%-Konfidenzgrenzen bei Darstellung der Bland-Altman-Plots in absoluten Zahlen nicht als parallel zur Abszisse verlaufende Geraden eingetragen werden. Bei prozentualer Werteauftragung der Ordinate wird jedoch die bestehende Abhängigkeit der Differenzen von der Spotanzahl aufgehoben und die Einzeichnung der Konfidenzgrenzen parallel zur Abszisse ermöglicht. (Vgl. 2.4.2).

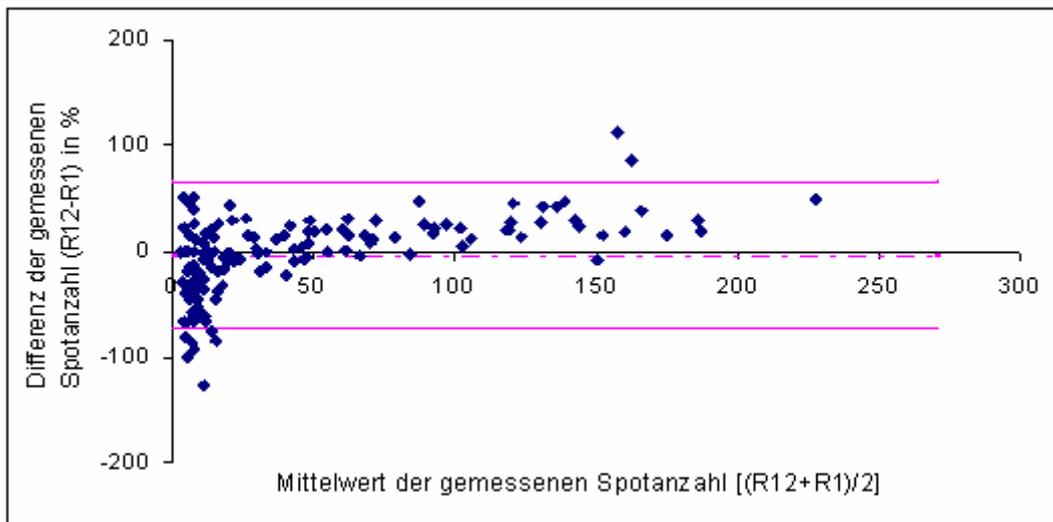
Das 95%-Konfidenzintervall gibt den Bereich an, in dem 95% aller Differenzwerte zu erwarten sind. Die Bestimmung der 95%-Konfidenzgrenzen erfolgte sowohl für die gesamte Spotanzahl, als auch für einzelne, eng gefasste Spotanzahlbereiche.

Die Bland-Altman-Plots in relativen Zahlen dienen weiterhin zur Visualisierung der Richtigkeit und Genauigkeit. Zur Veranschaulichung der Richtigkeit, als Maß für die Abweichung zwischen dem Mittelwert einzelner Messergebnisse und dem wahren Wert der Messgröße, wurde der Mittelwert der Differenzen als Parallele zur x-Achse eingezeichnet. Die 95%-Konfidenzintervalle dienen zur Beurteilung der Genauigkeit, als Maß dafür, wie nahe das Ergebnis am wahren Wert der Messgröße lag.

Das 95%-Vertrauensintervall für den Vergleich von R12 gegen R1 hatte die Intervallgrenzen: [-73.6%; 65.4%] (Abb. 24). Aus den Intervallgrenzen konnte

abgeleitet werden, dass R12 bis zu 73.6% weniger und bis zu 65.7% mehr Spots als R1 erkannte.

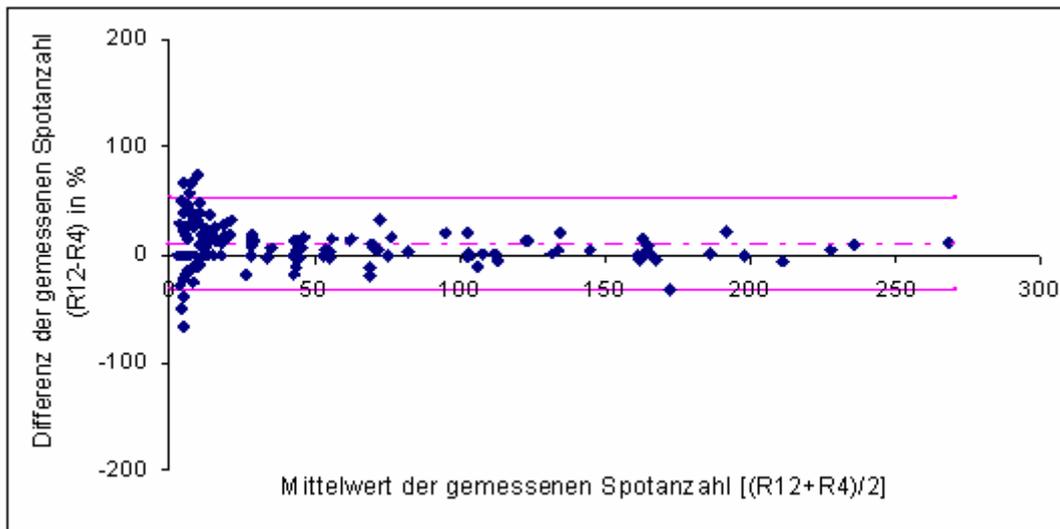
### KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Mausspots, prozentuale Darstellung



**Abb. 24:** Bland-Altman-Plot mit prozentualer Wertauftragung der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Mausspots. Die Konfidenzgrenzen [-76.6% und 65.4%], innerhalb deren 95% aller Differenzen erwartet werden können, entsprechen den parallel zur x-Achse verlaufenden Geraden. Sie werden aus dem Mittelwert der Differenzen (R12-R1) (dargestellt als gestrichelte Parallele zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R1) berechnet. Sie dienen weiterhin zur Visualisierung der Genauigkeit, während der Mittelwert der Differenzen die Richtigkeit veranschaulichte.

Die Bestimmung der 95%-Vertrauensintervalle für den Vergleich von R12 gegen R4 ergab die Grenzen [-32.5%; 52.9%] (Abb. 25). R12 erkannte demnach bis zu 32.5% weniger und maximal 52.9% mehr Spots als R4.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei  
Mausspots, prozentuale Darstellung**



**Abb. 25:** Bland-Altman-Plot mit prozentualer Werteauftragung der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (4 Bilder pro Well) bei Mausspots. Die Konfidenzgrenzen [-32.5% und 52.9%], innerhalb deren 95% aller Differenzen erwartet werden können, sind als parallel zur x-Achse verlaufende Geraden dargestellt. Sie ergeben sich aus dem Mittelwert der Differenzen (R12-R4) (entspricht der gestrichelten Parallelen zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R4). Sie dienen weiterhin zur Visualisierung der Genauigkeit, während der Mittelwert der Differenzen die Richtigkeit veranschaulichte.

Beim Vergleich der 95%-Intervallgrenzen (R12-R1) [-73.6%; 64.5%] und (R12-R4) [-32.5%; 52.9%] lag die obere Intervallgrenze (R12-R1) 11.6% über der Intervallgrenze (R12-R4). Bezüglich der unteren Konfidenzgrenzen konnten größere Unterschiede in den Werten beobachtet werden, die untere Grenze (R12-R1) betrug mehr als das Doppelte des (R12-R4) Wertes. Das 95%-Konfidenzintervall (R12-R4) fasste demnach einen deutlich engeren Bereich als das Intervall (R12-R1) ein, die Abweichungen zwischen R12 und R1 waren größer als die zwischen R12 und R4.

Die 95%-Konfidenzgrenzen (R12-R1) und (R12-R4) wurden für einzelne, kleinere Spotanzahlbereiche bestimmt (Tab. 13). Dies erlaubte präzisere Aussagen für konkrete Spotanzahlen zu treffen.

| Spotanzahl | 95%-Konfidenzintervalle<br>(R12-R1) | 95%-Konfidenzintervalle<br>(R12-R4) |
|------------|-------------------------------------|-------------------------------------|
| gesamt     | [-73.6%; 65.4%]                     | [-32.5%; 52.9%]                     |
| [3- 50]    | [-80.6%; 50.35%]                    | [-35.27%; 60.91%]                   |
| [50-100]   | [-8.81%; 41.28%]                    | [-17.61%; 29.9%]                    |
| [100-200]  | [-3.33%; 51.62%]                    | [-19.9%; 25.23%]                    |

**Tab. 13:** Übersicht der 95%-Konfidenzintervalle (in %) für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1) und 1.25-Zoom (R4) bei Mausspots. Innerhalb dieser Konfidenzgrenzen sind 95% aller Differenzen (R12-R1) bzw. (R12-R4) zu erwarten.

Im Spotanzahlbereich [3-50] lag die obere Grenze des 95%-Konfidenzintervalls (R12-R4) mit [60.91%] 10.6% über der oberen Grenze des (R12-R1)-Intervalls, die [50.35%] betrug. Bezüglich der unteren Konfidenzgrenze konnte in diesem Spotintervall die größte Abweichung aller 95%-Vertrauensintervalle gesehen werden: die untere Grenze des (R12-R1)-Intervalls zeigte mit [-80.6%] einen mehr als doppelt so hohen Wert wie die des (R12-R4)-Konfidenzintervalls mit [-35.27%] (Tab. 13). Bei Betrachtung des nächst höheren Spotanzahlbereiches [50-100] konnte für die obere Grenze des (R12-R4)-95%-Konfidenzintervalls [29.9%] ein um 11.4% niedrigerer Wert als für die obere Grenze des (R12-R1)-Intervalls [41.28%] nachgewiesen werden. Die untere Grenze des (R12-R4)-Vertrauensintervalls [-17.61%] differierte um 8.8% von der unteren Grenze des für (R12-R1) bestimmten 95%-Konfidenzintervalls [-8.81%] (Tab. 13).

Im Spotanzahlbereich [100-200] wiesen die oberen Grenzen der 95%-Konfidenzintervalle (R12-R4) und (R12-R1) die größten Unterschiede aller Intervalle auf: die obere Grenze des (R12-R1)-Intervalls [51.62%] hatte im

Vergleich zur oberen Grenze des (R12-R4)-Intervalls [25.23%] einen beinahe zweifach so hohen Wert. Zwischen den unteren Grenzen der 95%-Konfidenzintervalle (R12-R4) [-19.9%] und (R12-R1) [-3.33%] war eine Differenz von 16.6% nachzuweisen (Tab. 13).

### **3.6. Ergebnisse der Auswertung von Mausspots in Abhängigkeit von der Spotgröße**

Beim Vergleich der Systeme KS Elispot und KS Elispot compact wurden in der Messanordnung 1 Spots aller Durchmesser in die Messungen einbezogen und verschiedene Spotanzahlbereiche untersucht.

Das Ziel der Messanordnung 2 war, die Ergebnisse der Systeme abhängig von der Größe der Spots zu vergleichen und so zu prüfen, ob und wie die Spotgröße die Ergebnisse beeinflusste. Für Mausspots wurden die zwei Durchmesser Kategorien [20-100 $\mu$ m] und [100-300 $\mu$ m] gewählt.

#### **3.6.1. Mittelwerte der Differenzen für Mausspots mit Spotdurchmesser [20-100 $\mu$ m]**

Die Mittelwerte (MW) der Differenzen (R12-R1)<sub>[20-100 $\mu$ m]</sub> sowie (R12-R4)<sub>[20-100 $\mu$ m]</sub> wurden für die gesamte Spotanzahl, sowie für einzelne kleinere Spotanzahlbereiche berechnet. Die Ergebnisse wurden als absolute und relative Werte jeweils mit Standardabweichung (STD) dargestellt.

Für die gesamte Spotanzahl betrug der Mittelwert aller Differenzen (R12-R1)<sub>[20-100 $\mu$ m]</sub> [6.88 und 24.73%] und lag damit über dem Mittelwert aller Differenzen (R12-R4)<sub>[20-100 $\mu$ m]</sub> mit [2.66 und 14.26%] (Tab. 14, 15). Die Werte waren positiv, R12 zählte damit überwiegend mehr Spots als R1 und als R4. R1 erkannte weniger Spots als R4.

| <b>Spotanzahl n</b><br>[Mw(R12;R1);<br>Mw(R12; R4)] | <b>MW<br/>(R12-R1)</b><br>[20-100µm]<br>absolut | <b>STD</b> | <b>MW<br/>(R12-R4)</b><br>[20-100µm]<br>absolut | <b>STD</b> |
|---|---|------------|---|------------|
| n gesamt  | 6.88  | 18.34      | 2.66  | 8.51       |
| [0-50]  | 1.66  | 6.47       | 1.53  | 4.39       |
| [50-100]  | 49.45   | 21.28      | 9.52  | 15.87      |
| [100-200]   | 64.29   | 29.93      | 11.38   | 24.31      |

**Tab. 14:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) in absoluten Zahlen, ermittelt für Mausspots mit Spotdurchmessern [20-100µm].

| <b>Spotanzahl n</b><br>[Mw(R12;R1);<br>Mw(R12; R4)] | <b>MW<br/>(R12-R1)</b><br>[20-100µm]<br>relativ | <b>STD</b> | <b>MW<br/>(R12-R4)</b><br>[20-100µm]<br>relativ | <b>STD</b> |
|---|---|------------|---|------------|
| n <sub>≥</sub> 3                                    | 24.73%  | 0.4329     | 14.26%  | 0.3127     |
| [3-50]  | 14.29%  | 0.4026     | 15.15%  | 0.3462     |
| [50-100]  | 68.92%  | 0.2869     | 12.62%  | 0.2010     |
| [100-200]   | 55.81%  | 0.2459     | 9.95%   | 0.1701     |

**Tab. 15:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) in relativen Zahlen, ermittelt für Mausspots mit Spotdurchmessern [20-100µm].

In der Spotanzahlkategorie [3-50] betrug der Mittelwert aller Differenzen  $(R12-R1)_{[20-100\mu m]}$  [1.66 bzw. 14.29%]. Der Mittelwert aller Differenzen  $(R12-R4)_{[20-100\mu m]}$  hatte den Wert [1.53 bzw. 15.15%] (Tab. 14, 15). Demnach war keine nennenswerte Abweichung zwischen den Differenzmittelwerten  $(R12-R1)$  und  $(R12-R4)$  auszumachen. Beide Differenzmittelwerte waren positiv, folglich erkannte R12 überwiegend mehr Spots als R1 und als R4. Die Unterschiede zwischen R1 und R4 waren gering.

Bei einer Spotanzahl über 50 pro Well, konnten größere Abweichungen der Differenzmittelwerte nachgewiesen werden. Im Spotanzahlbereich [50-100] lag der Mittelwert der Differenzen  $(R12-R1)_{[20-100\mu m]}$  mit [49.45 bzw. 68.92%] etwa fünffach über dem Mittelwert der Differenzen  $(R12-R4)_{[20-100\mu m]}$ , der [9.52 bzw. 12.62%] betrug (Tab. 14, 15). Demnach identifizierte R12 mehrheitlich mehr Spots als R1 und als R4, wobei R1 in diesem Bereich weniger Spots als R4 erkannte.

In der Spotanzahlkategorie [100-200] konnten die größten Unterschiede in den Differenzmittelwerten beobachtet werden: der Mittelwert der Differenzen  $(R12-R1)_{[20-100\mu m]}$  war mit [64.29 bzw. 55.81%] deutlich höher als der Mittelwert der Differenzen  $(R12-R4)_{[20-100\mu m]}$  mit [11.38 bzw. 9.95%] (Tab. 14, 15). Die positiven Werte bedeuteten, dass R12 überwiegend mehr Spots als R1 bzw. als R4 zählte. Der deutlich höhere positive Wert des Differenzmittelwertes  $(R12-R1)$  drückte aus, dass R1 weniger Spots als R4 erkannte.

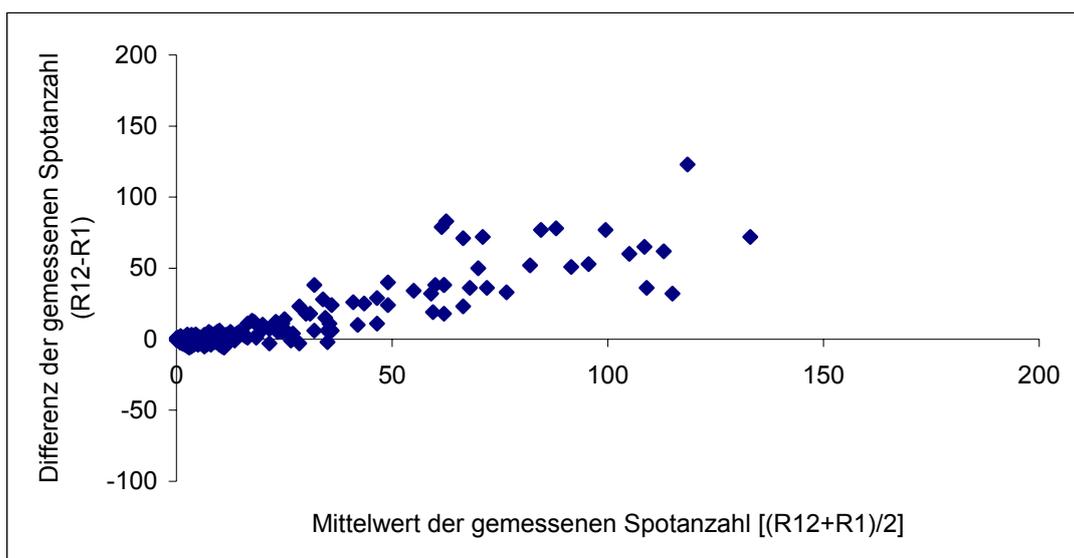
Beim Vergleich von R12 gegen R1 war mit zunehmender Spotanzahl pro Well ein Anstieg der absoluten und relativen Differenzmittelwerte  $(R12-R1)_{[20-100\mu m]}$  festzustellen. Je höher die Spotanzahl pro Well war, desto mehr differierten die von R12 und R1 gemessenen Werte, R1 maß dabei weniger Spots als R12: im Spotanzahlbereich [3-50] betrug der Differenzmittelwert  $(R12-R1)_{[20-100\mu m]}$  in relativen Zahlen [14.29%]. Er stieg in der Spotanzahlkategorie [50-100] auf den Wert [68.92%] an. Im Spotanzahlbereich [100-200] betrug er [55.81%] (Tab. 15). Die Differenzmittelwerte  $(R12-R1)_{[20-100\mu m]}$  in absoluten Zahlen nahmen mit zunehmender Spotanzahl pro Well ebenfalls immer größere Werte an: die Werte für die Spotanzahlkategorien [50-100] und [100-200] lagen mit [49.45] und [64.29] deutlich oberhalb des Differenzmittelwertes [1.66], der für den Spotbereich [0-50] galt (Tab. 14). Demnach zählte R1 mit zunehmender Spotzahl pro Well weniger Spots als R12.

Bezüglich der Mittelwerte der Differenzen  $(R12-R4)_{[20-100\mu\text{m}]}$  konnte ebenfalls eine Zunahme der absoluten Werte mit steigender Spotanzahl beobachtet werden: in den Spotanzahlkategorien [50-100] und [100-200] lagen die Mittelwerte der Differenzen mit [9.52] und [11.38] oberhalb des für den Spotbereich [3-50] geltenden Differenzmittelwertes [1.53]. Dieser Anstieg war deutlich kleiner, als der der Differenzmittelwerte  $(R12-R1)_{[20-100\mu\text{m}]}$  (Tab. 14). Die Mittelwerte der Differenzen  $(R12-R4)_{[20-100\mu\text{m}]}$  in relativen Zahlen zeigten bei steigender Spotanzahl keine Zunahme der Werte, in den Spotanzahlkategorien [50-100] und [100-200] lagen sie mit [12.62%] und [9.95%] geringfügig unterhalb des Differenzmittelwertes [15.15%], der für den Spotbereich [3-50] galt (Tab. 15).

### 3.6.2. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [20-100 $\mu\text{m}$ ] in absoluten Zahlen

Die Messdaten aus dem Vergleich R12 gegen R1 und R12 gegen R4 wurden mittels Bland-Altman-Plots in absoluten Zahlen graphisch dargestellt (Abb. 26, 27). In die Messungen gingen nur kleine Spots ein, deren Durchmesser im Bereich [20-100 $\mu\text{m}$ ] lagen.

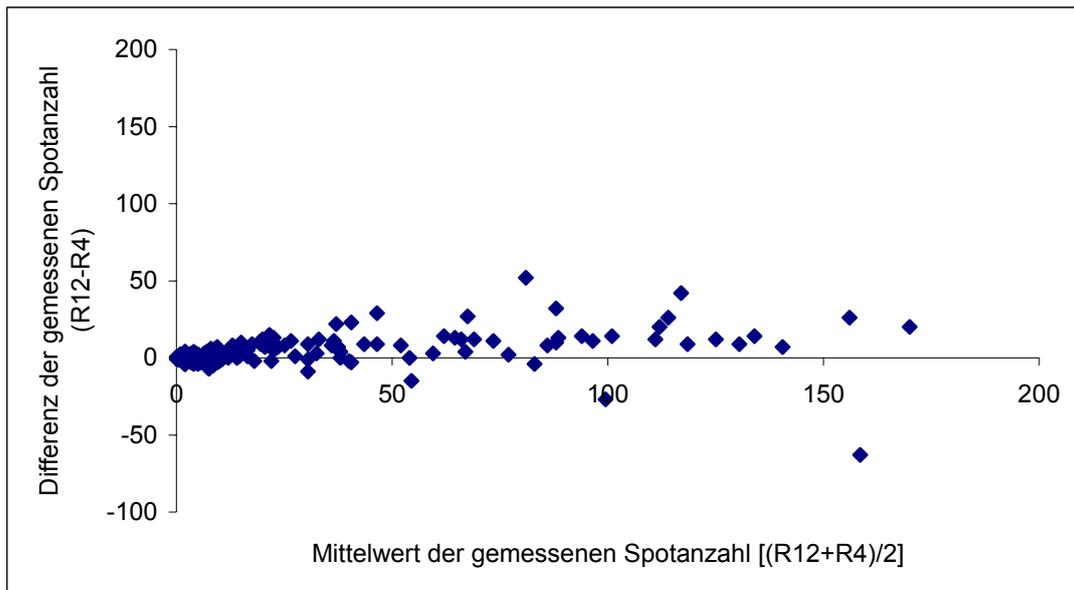
**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Mausspots mit Durchmessern [20-100 $\mu\text{m}$ ]**



**Abb. 26:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Mausspots

mit Spotdurchmessern [20-100µm]. Die Differenzwerte (R12-R1), dargestellt in Form von Rauten, waren mehrheitlich positiv, R12 zählte überwiegend mehr Spots als R1.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei  
Mausspots mit Durchmessern [20-100µm]**



**Abb. 27:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Mausspots mit Spotdurchmessern [20-100µm]. Die Differenzwerte, im Diagramm als Rauten veranschaulicht, waren mehrheitlich positiv, R12 zählte überwiegend mehr Spots als R4.

Im Spotanzahlbereich [0-50] bestanden im graphischen Verlauf von (R12-R1) und (R12-R4) geringe Abweichungen (Abb. 26, 27). Bei (R12-R1) lagen in dieser Kategorie 256 Differenzen im Wertebereich [-6; 40]. 150 dieser Differenzen waren positiv, 106 besaßen ein negatives Vorzeichen. Die Hälfte der negativen Differenzen hatte den Wert [-1]. Bei (R12-R4) wurden im entsprechenden Bereich 226 Differenzen gezählt, deren Werte sich über das Intervall [-9; 29] erstrecken. 160 der 226 Differenzen lagen im positiven Zahlenbereich. 66 der 226 Differenzen hatten Werte unter 0, wobei deutlich mehr als die Hälfte, 62%, den Differenzwert [-1] aufzeigte. Ein positiver Differenzmittelwert bedeutete, dass

R12 in diesem Well mehr Spots als R1 bzw. R4 erkannt hatte, während beim negativen Differenzwert die von R12 ermittelte Spotanzahl unter dem R1-bzw. R4-Ergebnis lag.

Der Anteil an positiven Differenzwerten im Spotanzahlbereich [0-50] lag demnach beim Vergleich R12 gegen R4 bei 70.8%, beim Vergleich R12 gegen R1 betrug er 58.6%. R12 maß somit überwiegend mehr Spots pro Well als R1 und als R4. Die positiven Differenzen (R12-R1) wiesen etwas höhere Werte als die Differenzen (R12-R4) auf, R4 zählte folglich mehr Spots pro Well als R1: 84.5% der positiven Differenzen (R12-R1) hatten einen Wert unter 10, 92.5% der positiven Differenzen (R12-R4) waren kleiner als 10, die restlichen Werte lagen darüber.

Im nächst höheren Spotanzahlbereich [50-100] nahmen die Divergenzen im Kurvenverlauf (R12-R1) und (R12-R4) zu (Abb. 26, 27): die 22 Differenzen (R12-R1) verteilten sich in dem Werteintervall [18; 83] und hatten ausschließlich positive Vorzeichen. Bei (R12-R4) befanden sich im entsprechenden Abschnitt 21 Differenzen im Intervall [-27; 52], davon 18 positive und 3 negative Differenzwerte. R12 maß demzufolge immer mehr Spots als R1 und fast in allen Wells mehr Spots als R4. Die positiven Differenzwerte (R12-R1) lagen höher als die positiven Differenzwerte (R12-R4). 16 der 18 positiven Differenzwerte (R12-R4), damit 89%, waren kleiner als 30. Bei (R12-R1) hatten 3 der 22 positiven Differenzen, damit 13.7%, Werte unter 30, die übrigen 19 lagen darüber. R1 zählte demnach weniger Spots als R4.

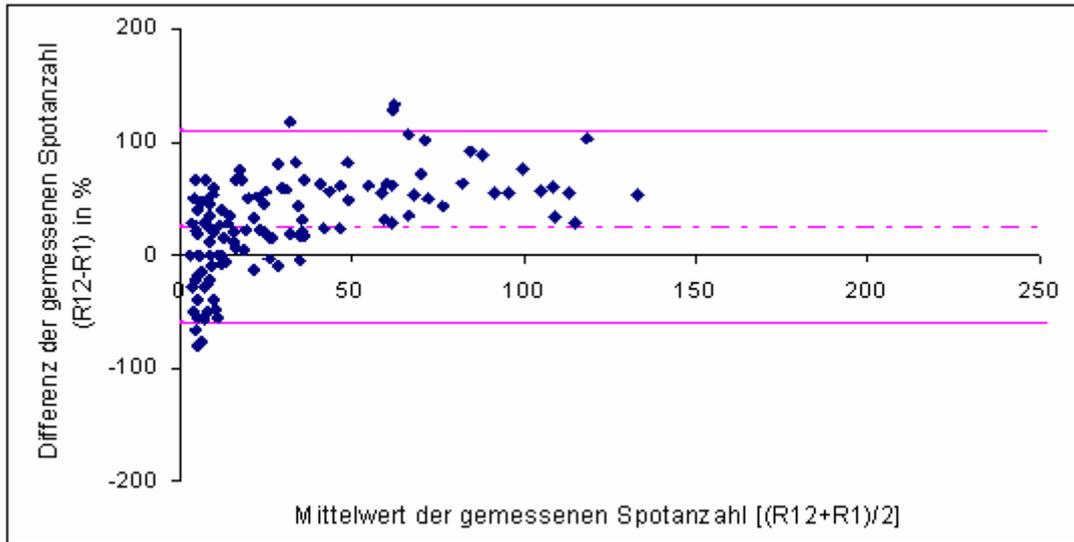
Im Spotanzahlbereich [100-200] wurden die deutlichsten Unterschiede in den Kurvenverläufen (R12-R1) und (R12-R4) beobachtet (Abb. 26, 27). Bei (R12-R1) lagen 7 Differenzwerte im Intervall [32; 123]. Alle Differenzwerte waren positiv, keine der Differenzen war kleiner als 30. Bei (R12-R4) wurden in diesem Spotanzahlbereich 13 Differenzen im Intervall [-63; 42] gezählt. Bis auf einen Wert [-63] hatten alle Differenzen positive Vorzeichen. 11 der 12 positiven Differenzwerte (R12-R4) waren kleiner als 30. Bei (R12-R1) war keine der Differenzen kleiner als 30. R12 erkannte somit fast immer mehr Spots als R1 und als R4. R1 zählte in diesem Bereich weit weniger Spots als R4.

Die aus den graphischen Verläufen in den Bland-Altman-Plots gewonnenen Beobachtungen stimmten mit den beim Vergleich der Differenzmittelwerte gewonnenen Ergebnissen (3. 2. 1) gut überein: bis zu einer Spotanzahl von 50 pro Well wurden in den graphischen Verläufen (R12-R1) und (R12-R4) geringere Unterschiede gesehen als in den Bereichen 50 bis 100 Spots und 100 bis 200 Spots pro Well.

### **3.6.3. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [20-100 $\mu$ m] in relativen Zahlen**

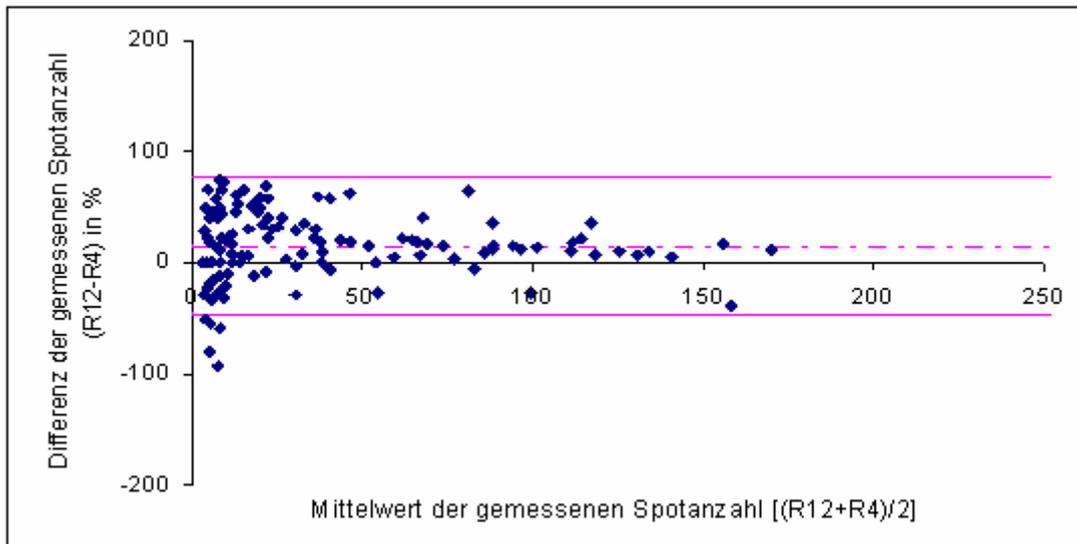
Die Messergebnisse aus dem Vergleich R12 gegen R1 und R12 gegen R4 wurden mittels der Bland-Altman-Plots auch als relative Werte graphisch veranschaulicht. Die 95%-Konfidenzgrenzen, innerhalb derer 95% der Differenzen  $(R12-R1)_{[20-100\mu m]}$  und  $(R12-R4)_{[20-100\mu m]}$  zu erwarten sind, wurden berechnet (Abb. 28, 29). Die prozentuale Werteauftragung auf der Ordinate wurde gewählt, um die bestehende Abhängigkeit der Differenzen von der Spotanzahl aufzuheben und die 95%-Konfidenzgrenzen parallel zur x-Achse einzeichnen zu können. Bei der Darstellung der Bland-Altman-Plots in absoluten Zahlen wäre die Darstellung der Konfidenzgrenzen als parallel zur x-Achse verlaufende Geraden aufgrund der mit der Spotanzahl ansteigenden Streuung der Differenzen nicht möglich gewesen.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei  
Mausspots mit Durchmessern [20-100µm], prozentuale Darstellung**



**Abb. 28:** Bland-Altman-Plot mit prozentualer Wertauftragung auf der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Mausspots mit Spotdurchmessern [20-100µm]. Die 95%-Konfidenzgrenzen [-60.1% und 109.6%] sind als zur x-Achse parallel verlaufende Geraden dargestellt. Sie ergeben sich aus dem Mittelwert der Differenzen (R12-R1) (entspricht der gestrichelten Parallelen zur Abszisse) +/- 1.96\* Standardabweichung (STD) der Differenzen (R12-R1).

**KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei Mausspots mit Durchmessern [20-100µm], prozentuale Darstellung**



**Abb. 29:** Bland-Altman-Plot mit prozentualer Wertauftragung auf der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Mausspots mit Spotdurchmessern [20-100µm]. Die 95%-Konfidenzgrenzen [-47% und 75.6%] entsprechen den parallel zur Abszisse verlaufenden Geraden. Sie werden aus dem Mittelwert der Differenzen (R12-R4) (dargestellt als gestrichelte Parallele zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R4) ermittelt.

Für den Vergleich der beiden Systeme ergab die Berechnung die folgenden 95%-Konfidenzintervalle:

$(R12-R1)_{[20-100\mu m]} : [-60.1\%; 109.6\%]$

$(R12-R4)_{[20-100\mu m]} : [-47\%; 75.6\%].$

Aus den Intervallgrenzen  $(R12-R1)_{[20-100\mu m]}$  war zu entnehmen, dass R12 maximal 109.6% mehr Spots als R1 und bis zu 60.1% weniger Spots als R1 erkannt hatte. Analog gab das 95%-Vertrauensintervall  $(R12-R4)_{[20-100\mu m]}$  an, dass R12 bis zu 75.6% mehr Spots und maximal 47% weniger Spots als R4 erkannt hatte. Die beiden Konfidenzintervalle differierten am stärksten in ihren oberen Grenzen: die Grenze des Intervalls  $(R12-R1)_{[20-100\mu m]}$  lag mit dem Wert

[109.6%] um 34% über der oberen Konfidenzgrenze (R12-R4)<sub>[20-100µm]</sub> mit [75.6%]. R1 maß demnach weniger Spots als R4.

Um genauere Aussagen für konkrete Spotanzahlen abzuleiten, wurden die 95%-Konfidenzgrenzen (R12-R1) und (R12-R4) für den Spotdurchmesser [20-100µm] auch für einzelne, engere Spotanzahlbereiche bestimmt.

| Spotanzahl | 95%-Konfidenzintervalle<br>(R12-R1)<br>[20-100µm] | 95%-Konfidenzintervalle<br>(R12-R4)<br>[20-100µm] |
|------------|---|---|
| gesamt     | [-60.1%; 109.6%]                                  | [-47.1%; 75.6%]                                   |
| [3- 50]    | [-66.23%; 94.82%]                                 | [-53.07%; 84.18%]                                 |
| [50-100]   | [11.54%; 126.3%]                                  | [-27.57%; 52.81%]                                 |
| [100-150]  | [6.64%; 104.99%]                                  | [-4.63%; 32.78%]                                  |

**Tab. 16:** 95%-Konfidenzgrenzen für den Vergleich KS Elispot (R12) und KS Elispot compact 0.65-Zoom (R1) bzw. 1.25-Zoom (R4) bei Mausspots mit Spotdurchmessern [20-100µm]. Innerhalb dieser Intervalle sind 95% der Differenzen (R12-R1) bzw. (R12-R4) zu erwarten. Die größten Unterschiede wurden bei Spotzahlen über 50 pro Well nachgewiesen.

Im Spotanzahlbereich [3-50] waren die geringsten Unterschiede in den 95%-Konfidenzgrenzen für den Vergleich von R12 gegen R1 und R12 gegen R4 zu erkennen (Tab. 16). Die obere Grenze des 95%-Konfidenzintervalls (R12-R1)<sub>[20-100µm]</sub> betrug [94.82%] und lag damit 10.6% über der oberen Grenze des 95%-Konfidenzintervalls (R12-R4)<sub>[20-100µm]</sub> [84.18%]. R12 zählte demnach in diesem Bereich bis zu 94.82% mehr Spots als R1 und bis zu 84.18% mehr Spots als R4. Bezüglich der unteren Grenzen lag der Wert des 95%-Konfidenzintervalls (R12-R1)<sub>[20-100µm]</sub> bei [-66.23%]. Die untere Grenze des 95%-Konfidenzintervalls (R12-R4)<sub>[20-100µm]</sub> lag mit [-53.07%] um 13.2% darunter. R12 maß demnach maximal 66.23% weniger Spots als R1 und maximal 53.07% weniger Spots als R4.

Im Spotanzahlbereich [50-100] zeigten beide Grenzen des 95%-Vertrauensintervalls  $(R12-R1)_{[20-100\mu\text{m}]}$  positive Werte (Tab. 16). Folglich maß R12 in diesem Bereich immer mehr Spots als R1. Die obere Grenze des 95%-Konfidenzintervalls  $(R12-R1)_{[20-100\mu\text{m}]}$  [126.3%] wies mehr als den zweifachen Wert der oberen Grenze des 95%-Vertrauensintervalls  $(R12-R4)_{[20-100\mu\text{m}]}$  [52.81%] auf, R1 zählte in diesem Bereich somit weniger Spots als R4. Die untere Grenze des 95%-Vertrauensintervalls  $(R12-R1)_{[20-100\mu\text{m}]}$  lag mit [11.54%] deutlich über der unteren Grenze des 95%-Vertrauensintervalls  $(R12-R4)_{[20-100\mu\text{m}]}$ , die [-27.57%] betrug.

Die größten Unterschiede in den oberen Intervallgrenzen waren im Spotanzahlbereich [100-150] zu sehen (Tab. 16). Die obere Grenze des 95%-Konfidenzintervalls  $(R12-R1)_{[20-100\mu\text{m}]}$  hatte mit [104.99%] einen mehr als dreifachen Wert der oberen Grenze des 95%-Vertrauensintervalls  $(R12-R4)_{[20-100\mu\text{m}]}$ , die [32.78%] betrug. Damit erkannte R1 in diesem Bereich wesentlich weniger Spots als R4. Bezüglich der unteren Intervallgrenzen konnten so große Abweichungen nicht beobachtet werden. Die untere Grenze des 95%-Konfidenzintervalls  $(R12-R1)_{[20-100\mu\text{m}]}$  lag bei [6.64%]. Die positive untere Intervallgrenze bedeutete, dass R12 in diesem Bereich stets mehr Spots zählte als R1. Der Wert der unteren Grenze des 95%-Vertrauensintervalls  $(R12-R4)_{[20-100\mu\text{m}]}$  betrug [-4.63%].

#### **3.6.4. Mittelwerte der Differenzen für Mausspots mit Spotdurchmesser [100-300 $\mu\text{m}$ ]**

Die Mittelwerte (MW) der Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  und  $(R12-R4)_{[100-300\mu\text{m}]}$  wurden für die gesamte Spotanzahl, sowie für einzelne Spotanzahlbereiche berechnet. Die Ergebnisse wurden als absolute und relative Werte, jeweils mit Standardabweichung (STD), dargestellt (Tab. 17, 18).

| Spotanzahl n                 | MW<br>(R12-R1)<br>[100-300µm]<br>absolut | STD   | MW<br>(R12-R4)<br>[100-300µm]<br>absolut | STD  |
|------------------------------|--|-------|--|------|
| [Mw(R12;R1);<br>Mw(R12; R4)] |  |       |  |      |
| n gesamt                     | 0.48                                     | 4.34  | 1.06                                     | 2.18 |
| [0-50]                       | 0.05                                     | 2.58  | 0.92                                     | 1.95 |
| [50-100]                     | 8.45                                     | 13.61 | 3.4                                      | 4.08 |

**Tab. 17:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) in absoluten Zahlen mit Standardabweichung (STD), ermittelt unter Einbeziehung von Mausspots mit Spotdurchmessern [100-300µm].

| Spotanzahl n                 | MW<br>(R12-R1)<br>[100-300µm]<br>relativ | STD    | MW<br>(R12-R4)<br>[100-300µm]<br>relativ | STD    |
|------------------------------|--|--------|--|--------|
| [Mw(R12;R1);<br>Mw(R12; R4)] |  |        |  |        |
| n <sub>≥3</sub>              | 1.09%                                    | 0.2430 | 9.4%                                     | 0.1671 |
| [3-50]                       | -0.69%                                   | 0.2380 | 9.99%                                    | 0.1760 |
| [50-100]                     | 12.11%                                   | 0.2368 | 5.01%                                    | 0.0611 |

**Tab. 18:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) in relativen Zahlen mit Standardabweichung (STD), ermittelt unter Einbeziehung von Mausspots mit Spotdurchmessern [100-300µm].

Für die Gesamtspotanzahl lag der Mittelwert aller Differenzen (R12-R1)<sub>[100-300µm]</sub> [absolut bzw. in %] mit [0.48 bzw. 1.09%] unter dem Mittelwert aller Differenzen

$(R12-R4)_{[100-300\mu\text{m}]}$  [1.06 bzw. 9.4%] (Tab. 17, 18). Die Mittelwerte waren positiv, demnach erkannte R12 überwiegend mehr Spots als R1 und R4.

Im Spotanzahlbereich [0-50] wies der Mittelwert aller Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  mit [0.05 bzw. -0.69%] einen geringeren Wert auf als der Mittelwert aller Differenzen  $(R12-R4)_{[100-300\mu\text{m}]}$ , der [0.92 bzw. 9.9%] betrug.

Im Spotanzahlbereich [50-100] konnte eine starke Zunahme des Mittelwertes der Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  beobachtet werden: im Vergleich zu dem Wert im Spotanzahlbereich [0-50] [0.05 bzw. -0.69%] betrug der Mittelwert der Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  im Bereich [50-100] mit [8.45 bzw. 12.11%] mehr als das Achtfache und lag damit deutlich höher als der entsprechende Mittelwert der Differenzen  $(R12-R4)_{[100-300\mu\text{m}]}$  [3.4 bzw. 5.01%] in diesem Bereich. R12 maß folglich mehr Spots pro Well als R1 und als R4. R1 zählte weniger Spots als R4.

### **3.6.5. Bland-Altman–Plots für Mausspots mit Spotdurchmesser [100-300 $\mu\text{m}$ ] in absoluten Zahlen**

In den Bland-Altman-Diagrammen gruppierten sich die Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  sowie  $(R12-R4)_{[100-300\mu\text{m}]}$  im Spotanzahlbereich [0-50] relativ dicht um die Abszisse (Abb. 30, 31). Die 197 Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  erstreckten sich über das Wertintervall [-5; 15], die 170 Differenzen  $(R12-R4)_{[100-300\mu\text{m}]}$  lagen in dem engeren Wertebereich [-4; 10]. Die Anzahl der negativen Differenzwerte  $(R12-R1)_{[100-300\mu\text{m}]}$  betrug 96; damit lag fast die Hälfte der Differenzen unterhalb der x-Achse. Diese negativen Differenzen nahmen maximal den Wert [-5] an, für über 60% konnte der Wert [-1] nachgewiesen werden. Bei  $(R12-R4)_{[100-300\mu\text{m}]}$  waren im Gegensatz dazu 36 der 170 Differenzwerte und damit 21% negativ. Über 70% der Werte entsprachen [-1], die restlichen hatten maximal den Wert [-4]. Bei  $(R12-R1)_{[100-300\mu\text{m}]}$  fanden sich unter allen 197 Differenzen 101 mit positivem Vorzeichen. 92% der Differenzen waren kleiner bzw. gleich [5]. Die 134 positiven Differenzen  $(R12-R4)_{[100-300\mu\text{m}]}$  hatten zu 97% Werte, die kleiner bzw. gleich [5] waren.

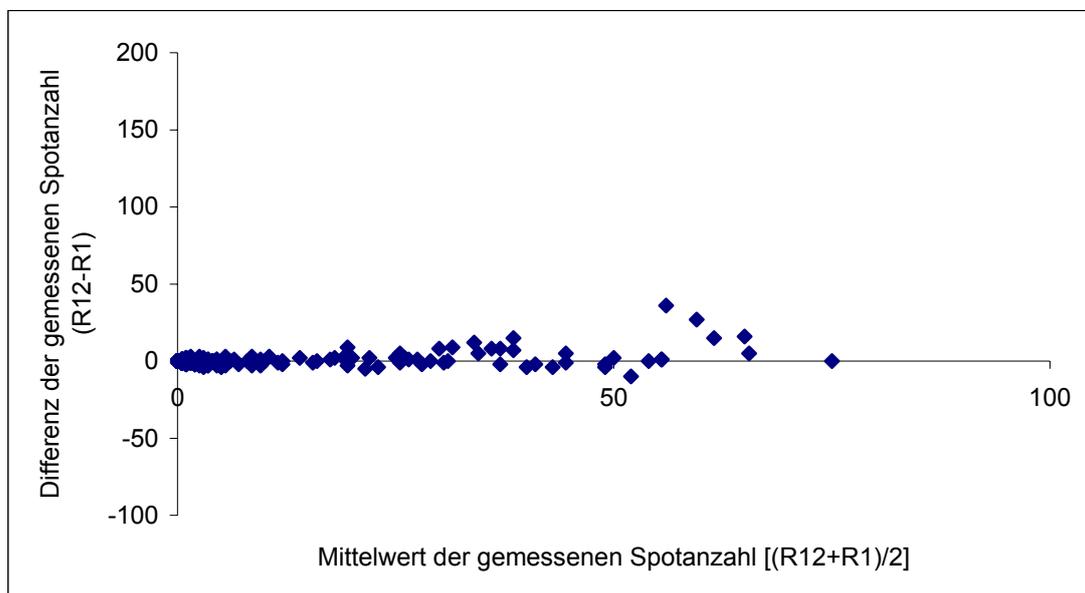
Im Spotanzahlbereich [50-100] zeigte der Verlauf von  $(R12-R1)_{[100-300\mu\text{m}]}$  durch seine größeren Differenzwerte im Positivbereich Abweichungen vom Verlauf  $(R12-R4)_{[100-300\mu\text{m}]}$  (Abb. 30, 31): das Wertintervall [-2; 10], innerhalb dessen die

10 Differenzen  $(R12-R4)_{[100-300\mu\text{m}]}$  lagen, war enger als das Intervall  $[-10; 36]$ , in dem sich die 11 Differenzen  $(R12-R1)_{[100-300\mu\text{m}]}$  verteilten.

Die 8 positiven Differenzen  $(R12-R4)_{[100-300\mu\text{m}]}$  wiesen als höchsten Wert [5] auf.

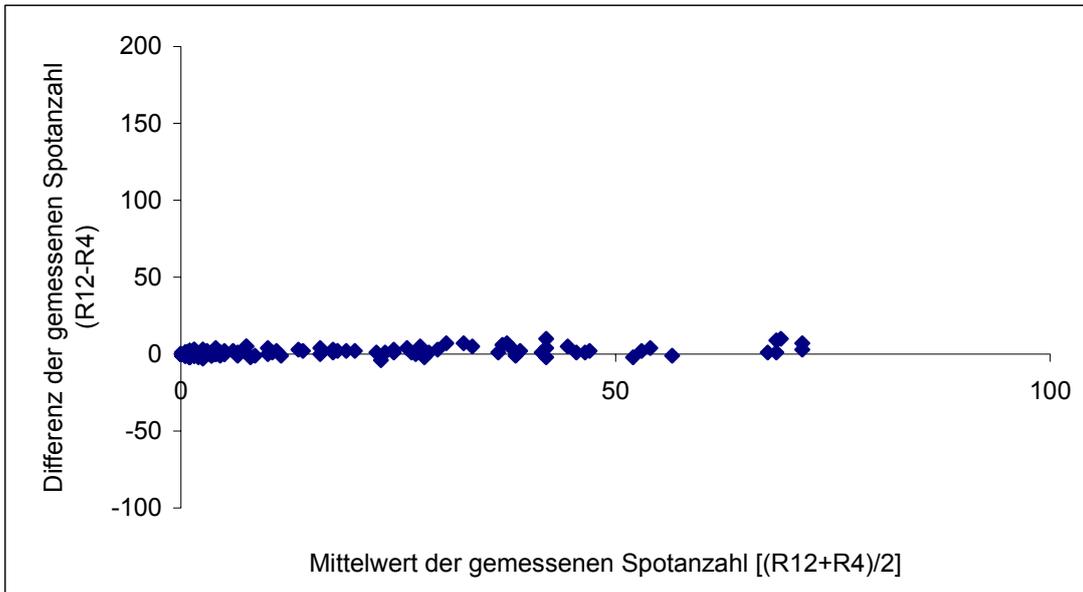
Bei  $(R12-R1)_{[100-300\mu\text{m}]}$  waren 6 der 10 positiven Differenzen kleiner bzw. gleich [5], die übrigen Werte waren in dem höheren Bereich [15-36] zu finden.

### KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Mausspots mit Durchmessern [100-300 $\mu\text{m}$ ]



**Abb. 30:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Mausspots mit Spotdurchmessern [100-300 $\mu\text{m}$ ]. Die Differenzwerte  $(R12-R1)$ , im Diagramm als Rauten dargestellt, waren klein und gruppierten sich dicht um die Abszisse.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei  
Mausspots mit Durchmessern [100-300µm]**



**Abb. 31:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Mausspots mit Spotdurchmessern [100-300µm]. Die Differenzwerte (R12-R4), im Diagramm als Rauten veranschaulicht, wiesen kleine Werte auf und gruppieren sich dicht um die Abszisse.

**3.6.6. Bland-Altman-Plots für Mausspots mit Spotdurchmesser [100-300µm] in relativen Zahlen**

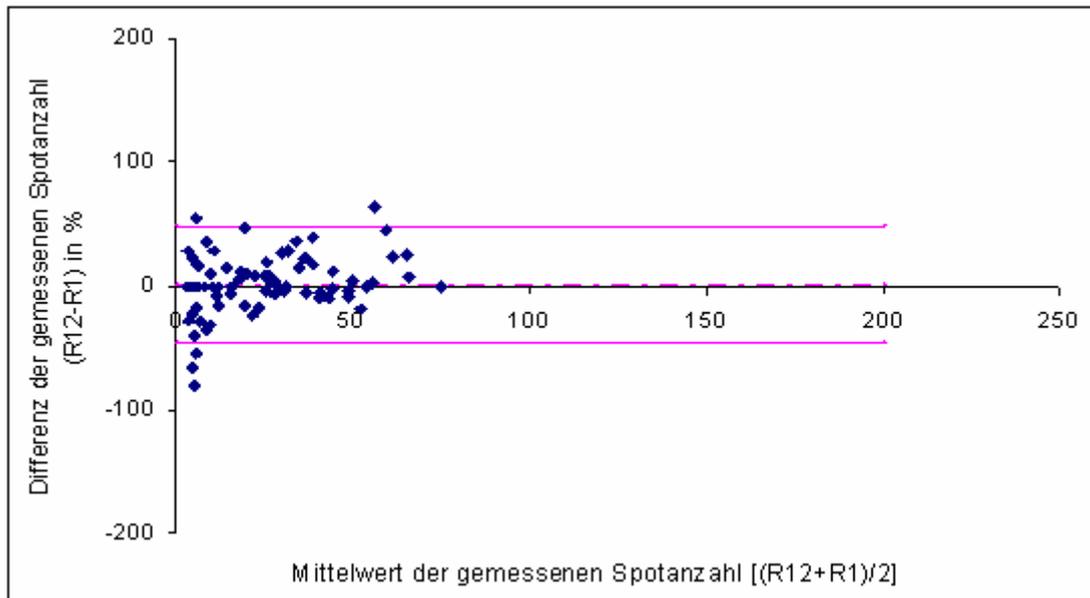
Die aus dem Vergleich von R12 gegen R1 und R12 gegen R4 gewonnenen Differenzdaten wurden mittels der Bland-Altman-Plots auch in relativen Zahlen veranschaulicht. Die Auftragung der Daten als relative Werte ermöglichte es, die 95%-Konfidenzgrenzen trotz der Unabhängigkeit der Differenzen von der Spotanzahl als zur Abszisse parallele Geraden einzuzichnen (Abb. 32, 33).

Die Berechnung der 95%-Konfidenzgrenzen ergab für den Vergleich von R12 gegen R1 und für den Vergleich von R12 gegen R4 die nachstehenden Intervalle:

$$(R12-R1)_{[100-300\mu m]}: [-46.5\%; 48.7\%]$$

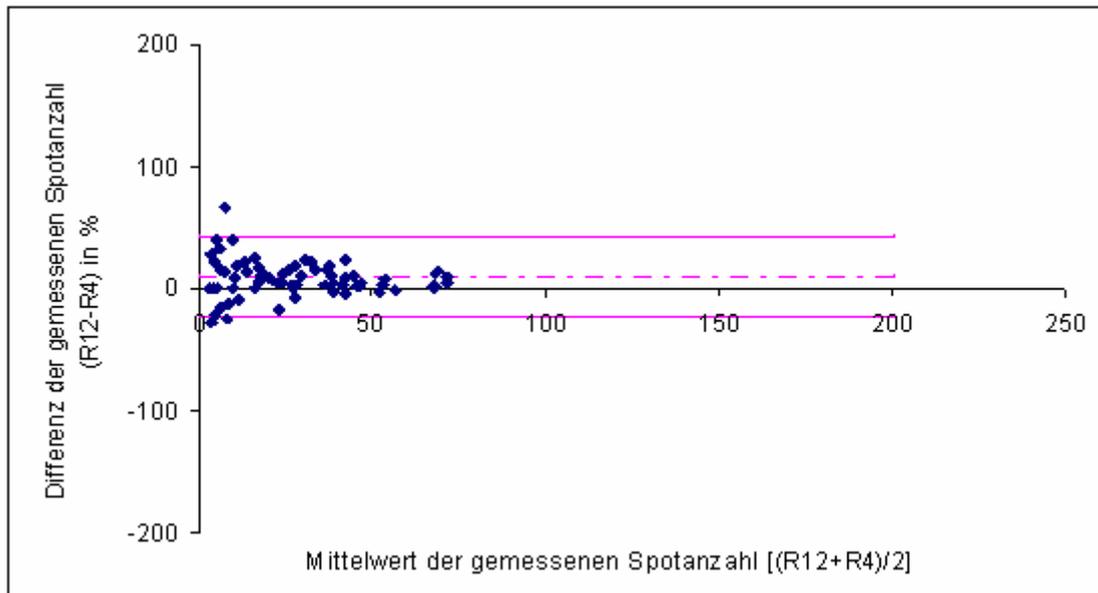
$$(R12-R4)_{[100-300\mu m]}: [-23.4\%; 42.2\%].$$

**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei  
Mausspots mit Durchmessern [100-300µm], prozentuale Darstellung**



**Abb. 32:** Bland-Altman-Plot mit prozentualer Werteauftragung auf der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Mausspots mit Spotdurchmessern [100-300µm]. Die Konfidenzgrenzen [-46.5% und 48.7%], innerhalb deren 95% aller Differenzen erwartet werden können, entsprechen den parallel zur x-Achse verlaufenden Geraden. Sie werden aus dem Mittelwert der Differenzen (R12-R1) (dargestellt als gestrichelte Parallele zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R1) ermittelt.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei  
Mausspots mit Durchmessern [100-300µm], prozentuale Darstellung**



**Abb. 33:** Bland-Altman-Plot mit prozentualer Werteauftragung auf der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Mausspots mit Spotdurchmessern [100-300µm]. Die 95%-Konfidenzgrenzen [-23.4% und 42.2%] sind als parallel zur Abszisse verlaufende Geraden veranschaulicht. Sie ergeben sich aus dem Mittelwert der Differenzen (R12-R4) (dargestellt als gestrichelte Parallele zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R4).

Innerhalb dieser Konfidenzgrenzen können 95% aller Differenzen erwartet werden. R12 hatte maximal 48.7% mehr Spots als R1 und bis zu 46.5% weniger Spots als R1 erkannt. Analog dazu hatte R12 maximal 42.2% mehr Spots als R4 und bis zu 23.4% weniger Spots erkannt.

Die beiden 95%-Vertrauensintervalle wiesen in ihren oberen Grenzen eine Differenz von 6.5% auf, wobei die obere Grenze des Intervalls  $(R12-R1)_{[100-300\mu m]}$  [48.7%] über der des Intervalls  $(R12-R4)_{[100-300\mu m]}$  [42.2%] lag. Die untere Grenze des Intervalls  $(R12-R1)_{[100-300\mu m]}$  war mit ihrem Wert [-46.5%] beinahe doppelt so hoch wie die untere Grenze des Intervalls  $(R12-R4)_{[100-300\mu m]}$ , die [-23.4%] betrug. Die 95%-Konfidenzintervalle  $(R12-R1)_{[100-300\mu m]}$  und  $(R12-R4)_{[100-300\mu m]}$  wurden zudem für zwei enge Spotanzahlbereiche bestimmt.

| Spotanzahl | 95%-Konfidenzintervalle<br>(R12-R1)<br>[100-300µm] | 95%-Konfidenzintervalle<br>(R12-R4)<br>[100-300µm] |
|------------|--|--|
| gesamt     | [-46.5%; 48.7%]                                    | [-23.4%; 42.2%]                                    |
| [3- 50]    | [-48.29%; 46.9%]                                   | [-25.21%; 45.2%]                                   |
| [50-100]   | [-33.55%; 61.68%]                                  | [-7.21%; 17.24%]                                   |

**Tab. 19:** 95%-Konfidenzintervalle (in %) für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1) bzw. gegen KS Elispot compact 1.25-Zoom (R4) bei Mausspots mit Spotdurchmessern [100-300µm]. Innerhalb dieser Intervalle sind 95% der Differenzen (R12-R1) bzw. (R12-R4) zu erwarten.

Im Spotanzahlbereich [3-50] unterschieden sich die beiden Konfidenzintervalle geringfügig in ihren oberen Intervallgrenzen. Die obere Grenze des  $(R12-R1)_{[100-300\mu m]}$ -95%-Vertrauensintervalls [46.9%] war um 1.7% höher als die obere Grenze des  $(R12-R4)_{[100-300\mu m]}$ -Vertrauensintervalls [45.2%]. Zwischen den unteren Grenzen der Intervalle war ein Unterschied von 23% nachzuweisen, wobei die untere Grenze des  $(R12-R1)_{[100-300\mu m]}$ -95%-Konfidenzintervalls den Wert [-48.29%] aufzeigte und die untere Grenze des Intervalls  $(R12-R4)_{[100-300\mu m]}$  [-25.21%] betrug.

Im Spotbereich [50-100] waren größere Abweichungen bezüglich der oberen und unteren Intervallgrenzen zu sehen. Die obere Grenze des 95%-Vertrauensintervalls  $(R12-R1)_{[100-300\mu m]}$  wies mit [61.68%] einen mehr als dreifach so hohen Wert im Vergleich zur oberen Grenze des Vertrauensintervalls  $(R12-R4)_{[100-300\mu m]}$  auf, für die der Wert [17.24%] ermittelt wurde. Die untere Grenze des  $(R12-R1)_{[100-300\mu m]}$ -95%-Vertrauensintervalls lag mit dem Wert [-33.55%] um mehr als das Vierfache über der unteren Grenze des Intervalls  $(R12-R4)_{[100-300\mu m]}$ , die [-7.21%] betrug.

### 3.6.7. Vergleich der 95%-Konfidenzgrenzen

Im folgenden wurden die, für alle Spotgrößen ermittelten 95%-Konfidenzintervalle den 95%-Konfidenzintervallen gegenübergestellt, die für Spots mit Durchmessern [20-100µm] und [100-300µm] berechnet wurden (Tab. 20).

| Spotdurchmesser  | 95%-Konfidenzintervalle<br>(R12-R1) | 95%-Konfidenzintervalle<br>(R12-R4) |
|------------------|-------------------------------------|-------------------------------------|
| gesamt: 20-300µm | [-73.6%; 65.4%]                     | [-32.5%; 52.9%]                     |
| 20-100µm         | [-60.1%; 109.6%]                    | [-47%; 75.6%]                       |
| 100-300µm        | [-46.5%; 48.7%]                     | [-23.4%; 42.2%]                     |

**Tab. 20:** Zusammenfassung der 95%-Konfidenzintervalle (in %) für den Vergleich KS Elispot (R12) mit KS Elispot compact 0.65-Zoom (R1) bzw. mit KS Elispot compact 1.25-Zoom (R4) bei Mausspots mit verschiedenen Spotdurchmessern. Die größten Unterschiede wurden bei der Messung kleiner Spots nachgewiesen.

Zunächst wurden die für alle Spots geltenden Konfidenzintervalle mit den für die Spotdurchmesser [20-100µm] bestimmten Intervallen verglichen.

Die 95%-Vertrauensintervalle für alle Spotgrößen hatten die folgenden Grenzen:

(R12-R1): [-73.6%; 65.4%]

(R12-R4): [-32.5%; 52.9%].

Die 95%-Vertrauensintervalle für Spots der Durchmesser [20-100µm] betragen:

(R12-R1)<sub>[20-100µm]</sub> : [-60.1%; 109.6%]

(R12-R4)<sub>[20-100µm]</sub> : [-47%; 75.6%].

Bei Gegenüberstellung der Intervalle für alle Spotgrößen und der für den Spotdurchmesserbereich [20-100µm] konnte folgendes festgestellt werden:

Bei (R12-R1) zeigte die obere Grenze des für die Spotdurchmesser [20-100µm] bestimmten 95%-Vertrauensintervalls mit [109.6%] einen um 44% höheren Wert als die für alle Spotdurchmesser geltende obere Grenze [65.4%]. Die untere Intervallgrenze (R12-R1) bei Spotgrößen [20-100µm] lag mit dem Wert [-60.1%] um 13.5% unter dem für alle Spotdurchmesser ermittelten unteren Intervallwert (R12-R1), der [-73.6%] betrug.

Bei (R12-R4) erweiterte sich das 95%-Vertrauensintervall an beiden Intervallgrenzen, nachdem nur Spots des Durchmessers [20-100µm] in die Messung einbezogen wurden: die obere Intervallgrenze erhöhte sich von [52.9%] um 23% auf den Wert [75.6%]. Die untere Grenze im Intervall (R12-R4)<sub>[20-100µm]</sub> verlagerte sich von [-32.5%] um 14.5% auf [-47%] weiter in den Negativbereich.

Zusammenfassend waren die 95%-Vertrauensintervalle für den Durchmesserbereich [20-100µm] sowohl beim Vergleich R12 gegen R1 als auch beim Vergleich R12 gegen R4 breiter als die für alle Spots geltenden Intervalle. Die größte Zunahme war bei der oberen Grenze des Intervalls (R12-R1)<sub>[20-100µm]</sub> nachzuweisen. Demnach waren die Abweichungen zwischen den Systemen bei kleineren Spotdurchmessern größer als bei Messungen von Spots aller Größen. Insbesondere beim Vergleich R12 gegen R1 maß R1 bei kleinen Spotdurchmessern weniger als R12. Zudem zählte R1 bei kleinen Spotdurchmessern weniger Spots als R4.

Analog zum durchgeführten Vergleich wurden den 95%-Vertrauensintervallen für alle Spotgrößen auch die für die Spotdurchmesser [100-300µm] bestimmten 95%-Vertrauensintervalle gegenübergestellt.

Die 95%-Konfidenzgrenzen für alle Spotgrößen betragen:

(R12-R1): [-73.6%; 65.4%]

(R12-R4): [-32.5%; 52.9%].

Die für die Spotdurchmesser [100-300µm] bestimmten 95%-Konfidenzgrenzen wiesen die folgenden Werte auf:

(R12-R1)<sub>[100-300µm]</sub>: [-46.5%; 48.7%]

(R12-R4)<sub>[100-300µm]</sub>: [-23.4%; 42.2%].

Die 95%-Vertrauensintervalle  $(R12-R1)_{[100-300\mu\text{m}]}$  sowie  $(R12-R4)_{[100-300\mu\text{m}]}$  ließen einen schmaleren Bereich mit enger zusammenliegenden oberen und unteren Intervallgrenzen erkennen als die für alle Durchmesser bestimmten 95%-Vertrauensintervalle.

Die obere Grenze des Intervalls  $(R12-R1)_{[100-300\mu\text{m}]}$  lag mit [48.7%] um 17% unter dem Wert [65.4%] der oberen Grenze des Vertrauensintervalls  $(R12-R4)$  für alle Spotdurchmesser. Die untere Grenze des 95%-Vertrauensintervalls  $(R12-R1)_{[100-300\mu\text{m}]}$  [-46.5%] lag um 27% höher als die untere Grenze des Vertrauensintervalls  $(R12-R1)$  [-73.6%] für alle Spotgrößen.

Bezüglich der Intervallgrenzen für den Vergleich von R12 gegen R4 fielen die Unterschiede geringer aus. Die obere Grenze des Vertrauensintervalls  $(R12-R4)_{[100-300\mu\text{m}]}$  hatte den Wert [42.2%]. Verglichen mit der oberen Grenze des für alle Spotdurchmesser geltenden Intervalls  $(R12-R4)$  [52.9%] lag sie 11% darunter. Die untere Grenze des Vertrauensintervalls  $(R12-R4)_{[100-300\mu\text{m}]}$  betrug [-23.4%] und damit 9% weniger als die untere Grenze des Vertrauensintervalls  $(R12-R4)$ , die bei [-32.5%] lag. Folglich differierten die Ergebnisse der beiden Systeme bei größeren Spots weniger als bei allen Spotgrößen, insbesondere beim Vergleich R12 gegen R1 verkleinerten sich die Unterschiede in den Messergebnissen.

Als weiterer Vergleich wurden die für die zwei Spotdurchmesser-kategorien [20-100 $\mu\text{m}$ ] und [100-300 $\mu\text{m}$ ] ermittelten 95%-Vertrauensintervalle einander gegenübergestellt.

Für die Spotdurchmesser [20-100 $\mu\text{m}$ ] galten die folgenden 95%-Konfidenzgrenzen:

$(R12-R1)_{[20-100\mu\text{m}]}$  : [-60.1%; 109.6%]

$(R12-R4)_{[20-100\mu\text{m}]}$  : [-47%; 75.6%].

Für die Spotdurchmesser [100-300 $\mu\text{m}$ ] wurden die folgenden 95%-Konfidenzgrenzen ermittelt:

$(R12-R1)_{[100-300\mu\text{m}]}$  : [-46.5%; 48.7%]

$(R12-R4)_{[100-300\mu\text{m}]}$  : [-23.4%; 42.2%].

Die 95%-Vertrauensintervalle für Spots des Durchmesserbereiches [20-100µm] waren breiter als die Intervalle für Spots des Durchmesserbereiches [100-300µm]. Folglich bestanden bei kleinen Spots mit Durchmessern [20-100µm] zwischen den beiden Systemen größere Unterschiede in den Messergebnissen als bei großen Spots mit Durchmessern [100-300µm].

Die oberen Grenzen der für kleine Spotdurchmesser geltenden Vertrauensintervalle  $(R12-R1)_{[20-100\mu m]}$  [109.6%] und  $(R12-R4)_{[20-100\mu m]}$  [75.6%] lagen deutlich über den für große Spots geltenden oberen Intervallsgrenzen  $(R12-R1)_{[100-300\mu m]}$  [48.7%] und  $(R12-R4)_{[100-300\mu m]}$  [42.2%]. Der Unterschied in den Werten betrug beim Vergleich R12 gegen R1 61%, bei Vergleich R12 gegen R4 33%.

Beim Vergleich R12 gegen R1 erkannte R12 bei kleinen Spots mit Durchmessern [20-100µm] demnach maximal bis zu 109% mehr Spots als R1, bei großen Spots mit Durchmessern [100-300µm] bis zu 48.7% mehr Spots als R1. Beim Vergleich R12 gegen R4 erkannte R12 bei kleinen Spots mit Durchmessern [20-100µm] bis zu 75.6% mehr Spots, bei großen Spots mit Durchmessern [100-300µm] bis zu 42.2% mehr Spots als R4. Somit erkannten R1 und R4 bei kleinen Spots deutlich weniger Spots als R12.

### ***3.7. Ergebnisse der Auswertung von Humanspots***

#### **3.7.1. Mittelwerte der Differenzen für Humanspots**

Entsprechend der Mausdatenauswertung wurden die Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) sowohl für die gesamte Spotanzahl als auch für einzelne Spotanzahlbereiche bestimmt und in absoluten Zahlen und als relative Werte angegeben (Tab. 21, 22).

| <b>Spotanzahl n</b><br>[MW(R12;R1);<br>MW(R12; R4)] | <b>MW</b><br><b>(R12-R1)</b><br>absolut | <b>STD</b> | <b>MW</b><br><b>(R12-R4)</b><br>absolut | <b>STD</b> |
|---|---|------------|---|------------|
| n gesamt  | 16.77                                   | 35.30      | 11.33                                   | 20.35      |
| [0-400]   | 11.70                                   | 20.76      | 10.37                                   | 18.61      |
| [0-200]   | 8.47                                    | 13.32      | 6.92                                    | 11.52      |
| [200-400]   | 54.27                                   | 44.53      | 59.54                                   | 29.30      |
| n <sub>≥</sub> 400                                  | 104.55                                  | 89.57      | 29.33                                   | 44.22      |
| [400-600]   | 61.33                                   | 75.90      | 20.50                                   | 51.58      |

**Tab. 21:** Mittelwerte (MW) der Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) in absoluten Zahlen für Humanspots.

| <b>Spotanzahl n</b><br>[MW(R12;R1);<br>MW(R12; R4)] | <b>MW</b><br><b>(R12-R1)</b><br>relativ | <b>STD</b> | <b>MW</b><br><b>(R12-R4)</b><br>relativ | <b>STD</b> |
|---|---|------------|---|------------|
| n <sub>≥</sub> 3                                    | 26%                                     | 0.2326     | 20.32%                                  | 0.2455     |
| [3-400]   | 26.66%                                  | 0.2367     | 21.38%                                  | 0.2490     |
| [3-200]   | 27.41%                                  | 0.2431     | 21.43%                                  | 0.2578     |
| [200-400]   | 18.6%                                   | 0.1345     | 20.85%                                  | 0.1158     |
| n <sub>≥</sub> 400                                  | 16.18%                                  | 0.1374     | 4.50%                                   | 0.0939     |
| [400-600]   | 11.68%                                  | 0.1504     | 3.92%                                   | 0.1169     |

**Tab. 22:** Mittelwerte (MW) aller Differenzen (R12-R1) und (R12-R4) mit Standardabweichung (STD) in relativen Zahlen für Humanspots.

Für die gesamte Spotanzahl betrug der Mittelwert aller Differenzen (MW) [absolut bzw. in %] für (R12-R1) [16.77 bzw. 26 %] und war damit höher als der Mittelwert aller Differenzen (R12-R4) [11.33 bzw. 20.32%]. Die Mittelwerte der Differenzen lagen im positiven Wertebereich. Es konnte somit angenommen werden, dass der überwiegende Teil der Differenzen Werte über Null aufwies und R12 damit mehrheitlich mehr Spots als R1 und R4 erkannt hatte. Der Mittelwert aller Differenzen (R12-R4) war kleiner als der aller Differenzen (R12-R1). R1 maß weniger Spots als R4.

Bei Betrachtung der einzelnen Spotanzahlbereiche wurde für das Intervall [0-400] festgestellt, dass der Mittelwert der Differenzen (R12-R1) mit [11.70 bzw. 26.66%] knapp über dem Mittelwert der Differenzen (R12-R4) mit [10.37 bzw. 21.38%] lag (Tab. 21, 22). Die positiven Mittelwerte der Differenzen zeigten, dass R12 überwiegend mehr Spots als R1 und R4 maß. R1 zählte im Intervall bis 400 Spots pro Well geringfügig weniger Spots als R4. Unterteilte man den Bereich bis 400 Spots pro Well weiter, so waren im Intervall [0-200] die Mittelwerte (R12-R1) mit [8.47 bzw. 27.41%] ebenfalls größer als die Mittelwerte (R12-R4) mit [6.92 bzw. 21.43%]. In der Spotanzahlkategorie [200-400] lagen die Mittelwerte (R12-R1) mit [54.27 bzw. 18.6%] unter den Mittelwerten (R12-R4) mit [59.54 bzw. 20.85%].

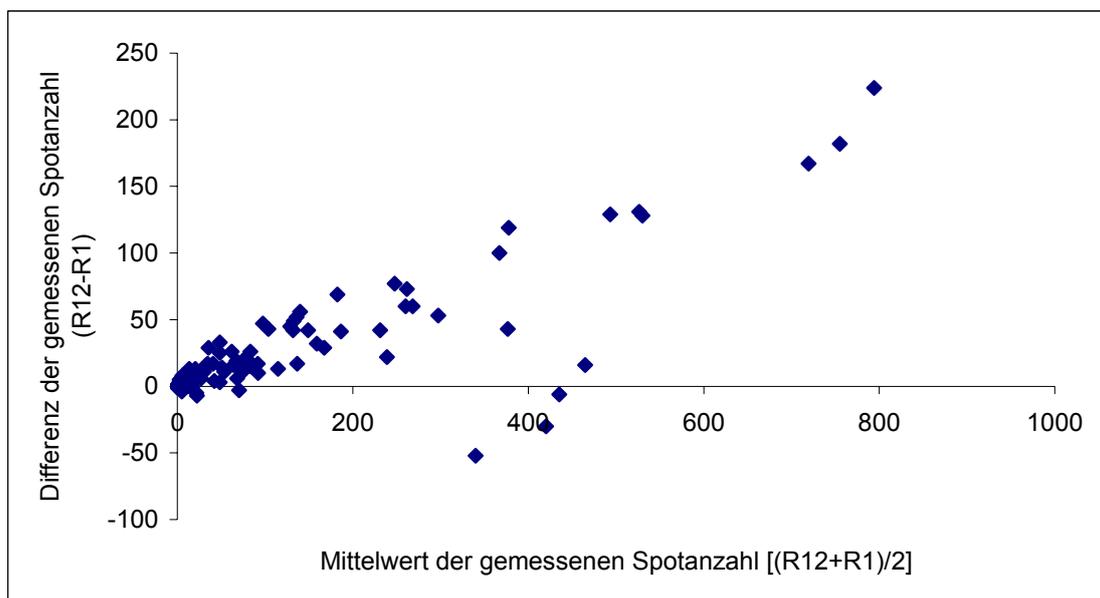
Im Vergleich zu den im Spotanzahlbereich unter 400 beschriebenen Abweichungen differierten die Mittelwerte der Differenzen (R12-R1) und (R12-R4) bei Spotzahlen über 400 pro Well deutlich stärker. Der Mittelwert der Differenzen (R12-R1) [104.55 bzw. 16.18%] betrug etwa das 3,5-fache des Differenzmittelwertes (R12-R4) [29.33 bzw. 4.5%]. Bei weiterer Aufteilung des Bereiches über 400 Spots pro Well in das Intervall [400-600] wurde beobachtet, dass der Mittelwert aller Differenzen (R12-R1) mit [61.33 bzw. 11.68%] etwa dreifach höher als der Mittelwert der Differenzen (R12-R4) mit [20.50 bzw. 3.92%] war. R1 zählte demnach bei Spotzahlen über 400 pro Well deutlich weniger Spots als R4. R12 maß überwiegend mehr Spots als R1 und als R4.

Die Tatsache, dass der Mittelwert aller Differenzen (R12-R1) über dem Mittelwert aller Differenzen (R12-R4) lag, konnte vorrangig auf diesen Anstieg der Differenzmittelwerte (R12-R1) bei Spotzahlen über 400 zurückgeführt werden, der bei den Differenzmittelwerten (R12-R4) nicht zu beobachten war.

### 3.7.2. Bland-Altman-Plots für Humanspots in absoluten Zahlen

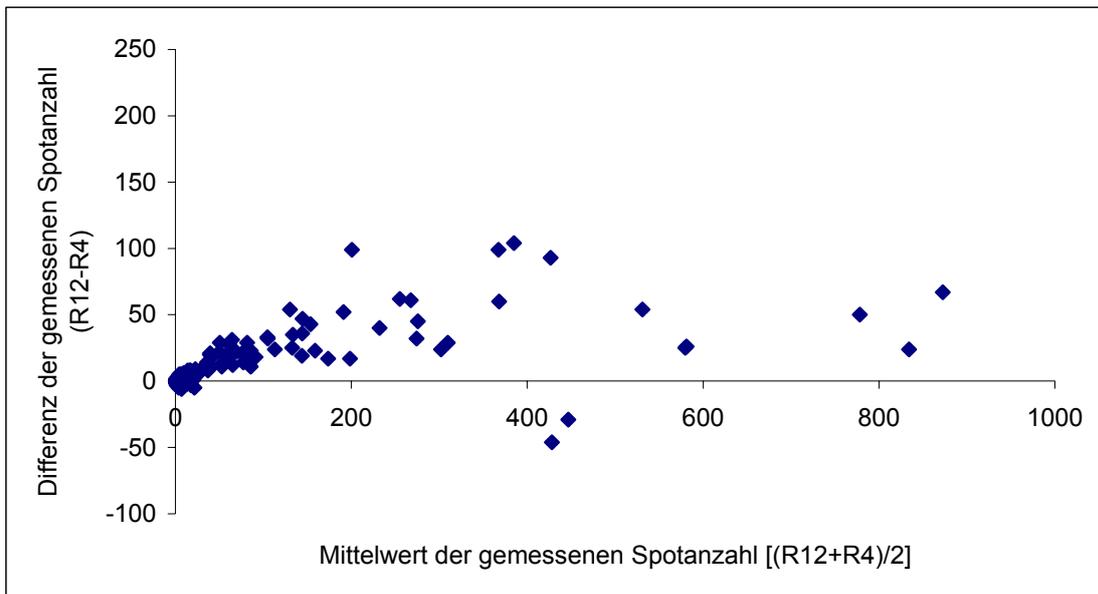
In den Bland-Altman-Plots wurden auf der Abszisse die Mittelwerte der von beiden Systemen pro Well gezählten Spotanzahl aufgetragen. Auf der Ordinate wurden die absoluten Differenzen zwischen der von beiden Systemen ermittelten Anzahl der Spots aufgezeichnet (Abb. 34, 35).

**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Humanspots**



**Abb. 34:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Humanspots. Die meisten Differenzen (R12-R1), im Diagramm als Rauten veranschaulicht, waren positiv. R12 zählte überwiegend mehr Spots als R1.

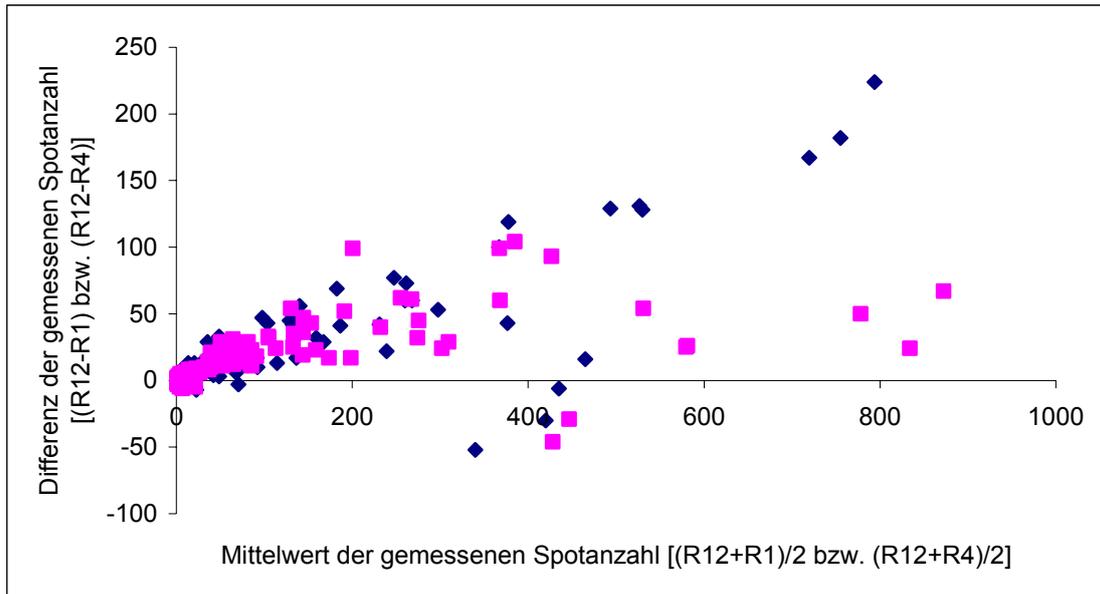
### KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei Humanspots



**Abb. 35:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Humanspots. Die Differenzwerte (R12-R4), im Diagramm als Rauten dargestellt, waren mehrheitlich positiv, R12 zählte überwiegend mehr Spots als R4.

Bei Gegenüberstellung der in den Bland-Altman-Plots in absoluten Zahlen aufgezeichneten Diagrammverläufe, wiesen die Systemvergleiche R12 gegen R1 und R12 gegen R4 in den Spotanzahlbereichen [0-200] bzw. [200-400] einen ähnlichen graphischen Verlauf auf, während sie bei Spotzahlen über 400 deutlich voneinander abwichen (Abb. 36).

### KS Elispot (R12) im Vergleich zu KS Elispot compact (R1, R4) bei Humanspots



**Abb. 36:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12, 12 Bilder pro Well) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) (veranschaulicht als Rauten  $\diamond$ ) sowie für KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) (veranschaulicht als Quadrate  $\square$ ) bei Humanspots. Die größten Unterschiede zwischen R1 und R4 wurden bei Spotzahlen über 400 Spots pro Well gesehen. Die meisten Differenzen waren positiv, R12 zählte überwiegend mehr Spots als R1 und R4.

Im Spotanzahlbereich [0-200] befanden sich bei (R12-R1) 145 Differenzwerte, die im Werteintervall [-7; 69] lagen. In den Bland-Altman-Plots entsprachen diese Differenzwerte den um die x-Achse verteilten Datenpunkten. 134 dieser Differenzen waren größer bzw. gleich 0. 91% der positiven Differenzen wiesen Werte unter 30 auf. 11 der Differenzen waren negativ. Ein positiver Differenzwert bedeutete, dass R12 in diesem Well mehr Spots als R1 gezählt hatte. Demnach erkannte R12 meist mehr Spots als R1. Bei (R12-R4) waren im Spotbereich [0-200] 157 Differenzwerte zu zählen. Sie erstreckten sich über das Werteintervall [-6; 54]. 120 der Differenzen lagen im Positivbereich. 92.5% der positiven Differenzwerte hatten einen Wert unter 30. 37 der Differenzen waren negativ. Folglich maß R12 überwiegend mehr Spots als R4.

In der Kategorie [200-400] Spots pro Well wurden bei (R12-R1) 10 Differenzwerte im Intervall [22; 119] nachgewiesen. Bei (R12-R4) fanden sich 11 Differenzen im Intervall [24; 104]. Alle Differenzwerte waren positiv. Damit zählte R12 immer mehr Spots pro Well als R1 und als R4.

Bei Spotzahlen über 400 Spots pro Well begannen die Kurvenverläufe zu divergieren: die auf der Ordinate aufgezeichneten Differenzwerte (R12-R1) stiegen im Positivbereich stark an und erreichten Werte, die deutlich über den Differenzwerten (R12-R4) lagen. Die Werte der Differenzen (R12-R4) zeigten diese Zunahme nicht: bei (R12-R1) wurden 9 Differenzen im Intervall [-30; 224] gezählt. 7 der 9 Differenzen wiesen positive Werte auf. Bei (R12-R4) lagen 11 Differenzen im Intervall [-46; 93]. 7 von ihnen waren positiv. Alle positiven Differenzwerte (R12-R4) waren kleiner als 100. Von den positiven Differenzen (R12-R1) lag eine unter 100, die übrigen 6 verteilten sich im Wertebereich 128 bis 224. R12 zählte damit mehrheitlich mehr Spots als R1 und als R4. R1 zählte bei Spotzahlen über 400 weniger Spots als R4.

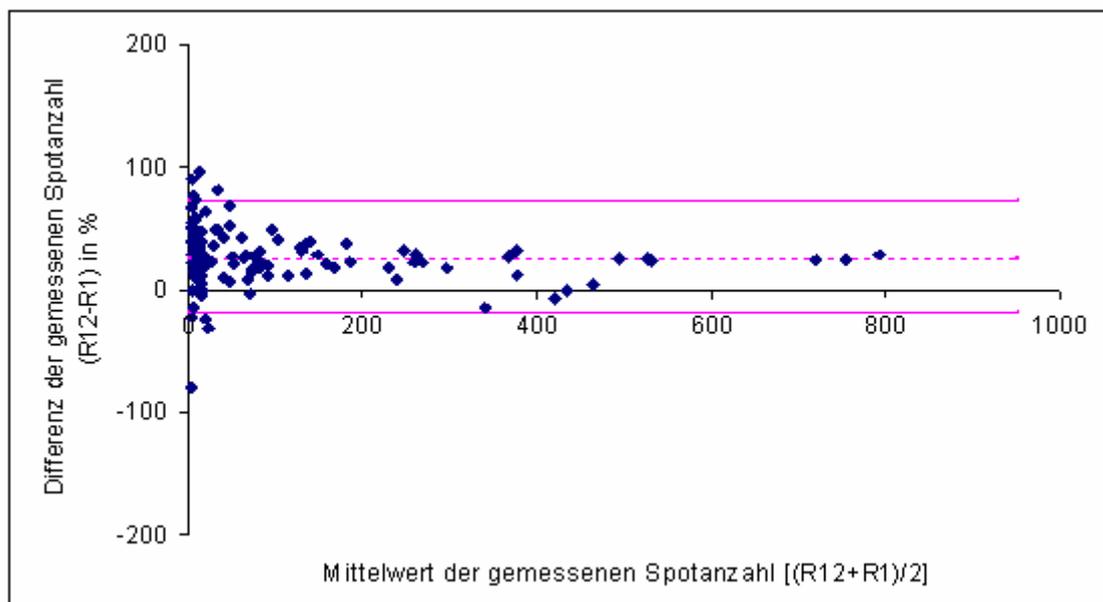
### **3.7.3. Bland-Altman-Plots für Humanspots in relativen Zahlen**

Neben den Bland-Altman-Plots in absoluten Zahlen wurden die Ergebnisse der Humanspotauswertung an beiden Systemen auch in relativen Zahlen mittels Bland-Altman-Plots zur Darstellung gebracht. Diese Veranschaulichung der Ergebnisse machte es möglich die Grenzen der 95%-Vertrauensintervalle zu bestimmen und als zur x-Achse parallel verlaufende Geraden einzuzeichnen. Aufgrund der Unabhängigkeit der Differenzen von der Spotanzahl wäre eine Auftragung der 95%-Konfidenzgrenzen parallel zur Abszisse bei Bland-Altman-Plots in absoluten Zahlen nicht möglich gewesen.

Die Berechnung der Konfidenzintervalle für die gesamte Spotanzahl ergab für den Vergleich R12 gegen R1 das Intervall [-19.2 %; 71.4 %], innerhalb dessen 95% der Differenzen erwartet werden konnten (Abb.40). Für den Vergleich R12 gegen R4 wurde das 95%-Konfidenzintervall [-27.3 %; 68.4 %] ermittelt (Abb. 41).

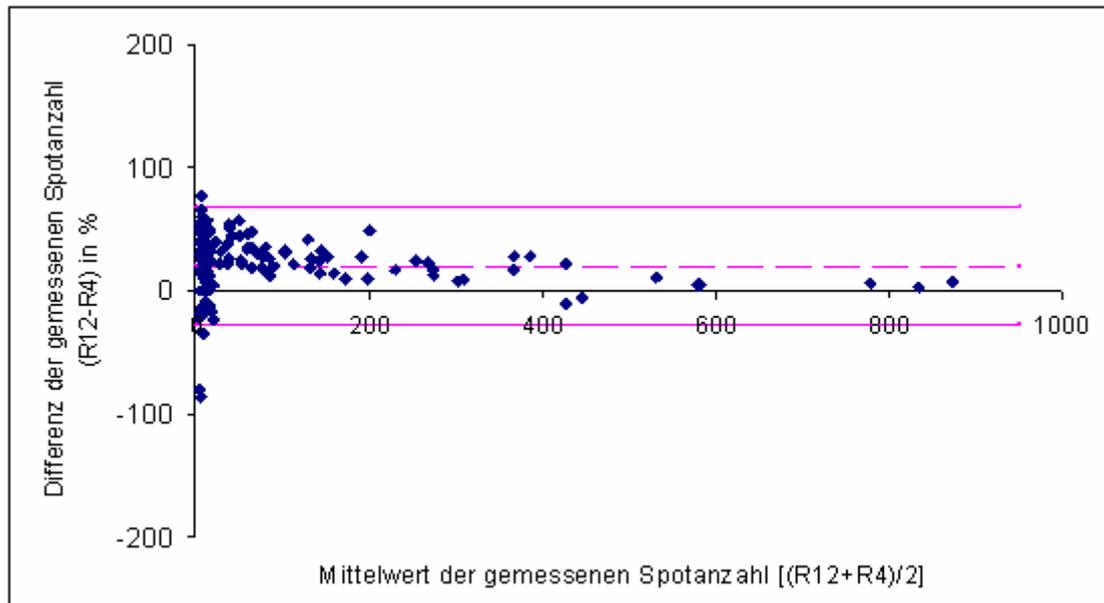
Die Grenzen der 95%-Vertrauensintervalle (R12-R1) [-19.2 %; 71.4 %] und (R12-R4) [-27.3 %; 68.4 %] schlossen einen ähnlich weiten Bereich ein. Die obere Grenze des Intervalls (R12-R1) lag um 3% höher als die obere Grenze des Intervalls (R12-R4). Bei den unteren Intervallgrenzen wurde für die Grenze (R12-R4) ein um 8% über der unteren Grenze des Intervalls (R12-R1) liegender Wert ermittelt.

### KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Humanspots, prozentuale Darstellung



**Abb. 40:** Bland-Altman Plot mit prozentualer Werteauftragung auf der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Humanspots. Die 95%-Konfidenzgrenzen [-19.2% und 71.4%], innerhalb deren 95% aller Differenzen erwartet werden können, sind als parallel zur Abszisse verlaufende Geraden dargestellt. Sie ergeben sich aus dem Mittelwert der Differenzen (R12-R1) (entspricht der gestrichelten Parallelen zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R1). Sie dienen weiterhin zur Visualisierung der Genauigkeit, während der Mittelwert der Differenzen die Richtigkeit veranschaulichte.

### KS Elispot (R12) im Vergleich zu KS Elispot compact 1.25 (R4) bei Humanspots, prozentuale Darstellung



**Abb. 41:** Bland-Altman-Plot mit prozentualer Wertauftragung auf der Ordinate für den Vergleich KS Elispot (R12) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) bei Humanspots. Die 95%-Konfidenzgrenzen [-27.3% und 68.4%] sind als zur Abszisse parallel verlaufende Geraden eingezeichnet. Sie ergeben sich aus dem Mittelwert der Differenzen (R12-R4) (entspricht der gestrichelten Parallelen zur Abszisse)  $\pm 1.96 \cdot$  Standardabweichung (STD) der Differenzen (R12-R4).

Die 95%-Konfidenzintervalle wurden darüber hinaus für einzelne, eng gefasste Spotanzahlbereiche bestimmt, um so auch Informationen für konkrete Spotanzahlen zu erhalten.

| Spotanzahl | 95%-Konfidenzintervalle<br>(R12-R1) | 95%-Konfidenzintervalle<br>(R12-R4) |
|------------|-------------------------------------|-------------------------------------|
| gesamt     | [-19.2%; 71.4%]                     | [-27.3%; 68.4%]                     |
| 3- 50      | [-26.44%; 83.24%]                   | [-38.48%; 77.44%]                   |
| 50-100     | [-1.97%; 45.43%]                    | [6.13%; 54.94%]                     |
| 100-200    | [7.09%; 50.27%]                     | [4.37%; 42.61%]                     |
| 200-400    | [-8.3%; 45.5%]                      | [-2.31%; 44.01%]                    |
| 400-600    | [-18.41%; 41.77%]                   | [-19.46%; 27.3%]                    |
| 600-900    | [19.84%; 30.51%]                    | [0.68%; 10.64%]                     |

**Tab. 23:** 95%-Konfidenzintervalle (in %) für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1) bzw. gegen KS Elispot compact 1.25-Zoom (R4) bei Humanspots. Die größten Unterschiede wurden bei Spotzahlen über 400 pro Well beobachtet.

Beim Vergleich der Vertrauensintervalle (R12-R1) und (R12-R4) konnten die größten Unterschiede ab 400 Spots pro Well nachgewiesen werden (Tab. 23).

Im Bereich von 400 bis 600 Spots pro Well lag die obere Grenze des Intervalls (R12-R1) mit [41.77%] 14.5% über der oberen Grenze des Intervalls (R12-R4) mit [27.3%]. Die unteren Intervallgrenzen unterschieden sich kaum. Die größte Abweichung wurde bei Spotzahlen über 600 pro Well beobachtet. Die obere Grenze des 95%-Vertrauensintervalls (R12-R1) war mit [30.51%] um 19.9% höher als die des Intervalls (R12-R4) mit [10.64%]. Die untere Grenze des Vertrauensintervalls (R12-R1) [19.84%] lag um ein Vielfaches über dem unteren Wert von (R12-R4) [0.68%]. Bei Spotzahlen über 400 pro Well erkannte R1 demnach weniger Spots als R4.

Die Grenzen der 95%-Intervalle (R12-R1) und (R12-R4) zeigten in den 4 Spotanzahlbereichen unter 400 Spots pro Well wesentlich kleinere Abweichungen als die oben beschriebenen Intervallgrenzen bei Spotzahlen über 400 pro Well.

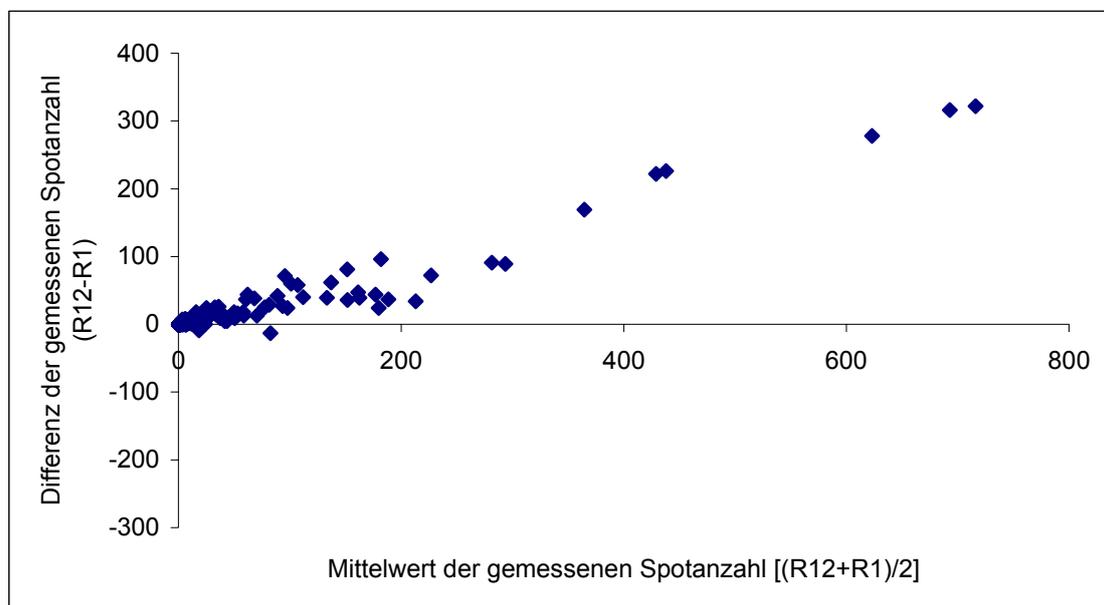
### 3.8. Ergebnisse der Auswertung von Humanspots in Abhängigkeit von der Spotgröße

#### 3.8.1. Bland-Altman-Plots in absoluten Zahlen

In dieser Messanordnung (Messanordnung 2) sollte analog zu Mausspots auch bei Humanspots der Einfluss der Spotgrößen auf die Ergebnisse der Systeme geprüft werden. Die Humanplatten enthielten überwiegend kleine Spots. Um die Unterschiede der Messsysteme in den kleineren Größenbereichen klarer herauszuarbeiten, wurden 3 Größenkategorien gewählt: [20-50µm], [50-100µm], [100-300µm].

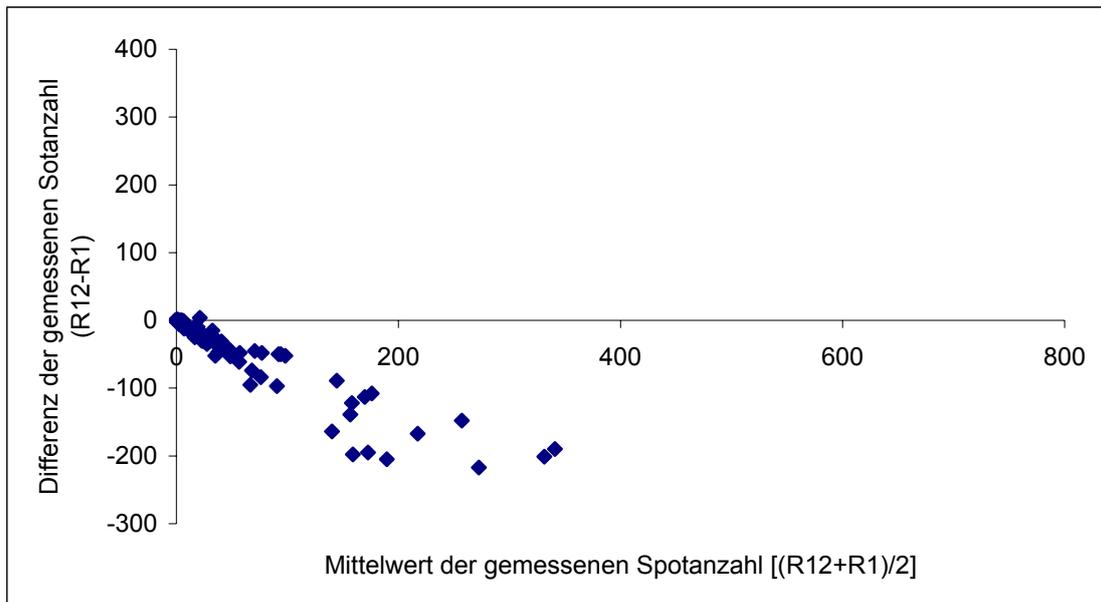
Die gewonnenen Auswertungsdaten aus dem Vergleich R12 gegen R1 und R12 gegen R4 in den genannten drei Größenbereichen wurden mittels der Bland-Altman-Plots graphisch veranschaulicht. Für den Vergleich R12 gegen R1 entstanden demnach die nachstehenden drei Bland-Altman-Plots in absoluten Zahlen (Abb. 42, 43, 44).

**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei Humanspots mit Durchmessern [20-50µm]**



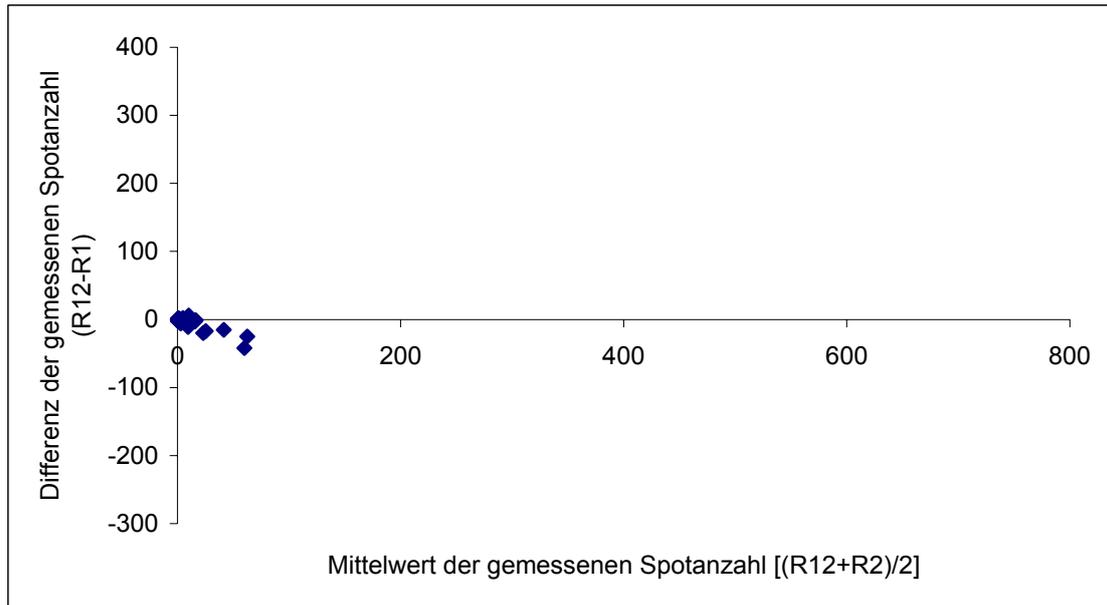
**Abb. 42:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Humanspots mit Spotdurchmessern [20-50µm]. Auffällig waren die mit zunehmender Spotzahl größer werdenden Differenzen (R12-R1).

**KS Elispot (R12) im Vergleich zu KS Elispot 0.65 (R1) bei Humanspots mit Durchmessern [50-100µm]**



**Abb. 43:** Bland-Altman-Plot in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Humanspots mit Spotdurchmessern [50-100µm]. Auffällig waren die mit zunehmender Spotzahl immer größer werdenden negativen Differenzen.

**KS Elispot (R12) im Vergleich zu KS Elispot compact 0.65 (R1) bei  
Humanspots mit Durchmessern [100-300µm]**



**Abb. 44:** Bland-Altman-Plots in absoluten Zahlen für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) bei Humanspots mit Spotdurchmessern [100-300µm]. Auffällig waren die mit zunehmender Spotzahl immer größeren negativen Differenzen.

Die Bland-Altman-Plots für Spots mit Durchmessern [20-50µm] (Abb. 42) zeigten, dass die Differenzen (R12-R1) mit zunehmender Spotanzahl immer größere Werte annahmen und in ihrem Verlauf das Bild einer oberhalb der x-Achse linear ansteigenden Geraden ergaben. Die zunehmenden positiven Differenzwerte bedeuteten, dass R1 mit steigender Spotanzahl deutlich weniger Spots als R12 erkannte.

Die für Spots der Durchmesser [50-100µm] erstellten Bland-Altman-Plots (Abb. 43) differierten von diesem Verlauf. Die Differenzwerte (R12-R1) wiesen fast ausschließlich negative Werte auf, die mit steigender Spotanzahl immer weiter in den Negativbereich reichten. Sie ergaben auch das Bild einer linear steigenden Geraden, die in diesem Fall unterhalb der Abszisse lag. Die ansteigenden negativen Differenzen drückten aus, dass R1 mit zunehmender Spotanzahl mehr Spots zählte als R12.

In den Bland-Altman-Plots für Spots mit Durchmessern [100-300µm] (Abb. 44) lagen weit weniger Differenzwerte als in den anderen zwei Größenbereichen [20-50µm] und [50-100µm], der graphische Verlauf entsprach aber dem schon im Größenbereich [50-100µm] beschriebenen Verlauf: die Differenzen hatten zum größten Teil negative Werte, die mit zunehmender Spotanzahl weiter anstiegen und in ihrem Verlauf mit einer unterhalb der x-Achse linear verlaufenden Gerade vereinbar waren.

Die Bestimmung der relativen Größenanteile in den drei Durchmesser kategorien unter Verwendung der am System KS Elispot (R12) gewonnen Daten hatte die folgende Verteilung ergeben: im Größenbereich [20-50µm] lagen 76.23% aller Humanspots, im Größenbereich [50-100µm] wurden 21.39% aller Spots nachgewiesen, im Größenbereich [100-300µm] befanden sich 2.28% der Spots. Die Verteilung der Spots auf die drei Durchmesserbereiche wurde nun auch anhand der am System KS Elispot compact 0.65-Zoom (R1) für die jeweilige Kategorie gezählten Spots bestimmt und mit der am System R12 ermittelten verglichen (Tab. 24).

| Spotdurchmesser<br>in µm | Anteil der Spots in %<br>KS Elispot<br>(R12) | Anteil der Spots in %<br>KS Elispot compact 0.65<br>(R1) |
|--------------------------|--|--|
| [20-50µm]                | 76.23%                                       | 47.8%  |
| [50-100µm]               | 21.39%                                       | 48.79%   |
| [100-300µm]              | 2.28%  | 3.4%   |

**Tab. 24:** Vergleich der am System KS Elispot (R12) und KS Elispot compact 0.65-Zoom (R1) ermittelten Verteilung der Humanspots in Abhängigkeit vom Durchmesser. Der R12-Wert für den relativen Spotanteil im Durchmesserbereich [20-100µm] betrug das ca. 1.5-fache des R1-Wertes. Im Bereich [50-100µm] und [100-300µm] lag der R1-Wert ca. zweifach über dem am R12 ermittelten Spotanteil.

Es wurde nachgewiesen, dass die aus R1-Daten berechneten relativen Anteile der Spots an den drei Durchmesserbereichen nicht mit den aus R12-Daten gewonnenen Anteilen übereinstimmten (Tab. 24): bei R1 betrug der Anteil der Spots im Durchmesserbereich [20-50 $\mu$ m] 47.8%, im Größenintervall [50-100 $\mu$ m] lag er bei 48.79% und im Durchmesserbereich [100-300 $\mu$ m] wies er den Wert 3.4% auf.

Der R12-Wert [76.23%] für den relativen Spotanteil im Durchmesserbereich [20-100 $\mu$ m] lag ca. 1.6-fach über dem R1-Wert [47.8%]. Für den Durchmesserbereich [50-100 $\mu$ m] lieferte R1 mit [48.79%] einen mehr als doppelt so hohen Spotanteil als R12 mit [21.39%]. Bei Durchmessern [100-300 $\mu$ m] konnte die geringste Abweichung beobachtet werden, der R1-Wert [3.4%] für diesen Bereich lag ca. 1.5-fach über dem R12-Wert [2.28%].

Die von R12 und R1 ermittelten Verteilungen der Humanspots auf die drei Durchmesser kategorien differierten deutlich. Die graphischen Verläufe in den Bland-Altman-Plots konnten auf diese Abweichungen zurückgeführt werden: Im Bereich [20-50 $\mu$ m] zählte R12 viel mehr Spots als R1, da R1 fast 30% weniger Spots als R12 in diesen Größenbereich eingeordnet hatte. Bei Durchmessern [50-100 $\mu$ m] und [100-300 $\mu$ m] lagen die R1-Ergebnisse über den R12 Ergebnissen, R12 zählte hier also viel weniger Spots als R1, da R1 diesen Größenbereichen einen deutlich höheren Spotanteil zugerechnet hatte als R12.

Das Auftreten der Abweichungen in den am System KS Elispot und am System KS Elispot compact 0.65-Zoom erhobenen Daten lieferte wichtige Informationen für den Vergleich der Systeme und wird unter 4.2.3. erörtert: R1 hatte einen Großteil der kleinen Spots, die R12 noch dem Durchmesserbereich [20-50 $\mu$ m] zugeordnet hatte, schon im nächst höheren Bereich lokalisiert. Ebenso hat R1 einen Teil der Spots, die R12 noch im Bereich [50-100 $\mu$ m] lokalisierte, bereits dem nächst höheren Intervall [100-300 $\mu$ m] zugerechnet und damit die gesamte Verteilung der Größenanteile verschoben. Als Ursache dieses Phänomens konnte die unterschiedliche Auflösung mit der die beiden Systeme arbeiten, angesehen werden.

### **3.8.2. Bland-Altman-Plots in relativen Zahlen**

Die weitere Auswertung der in der durchmesserabhängigen Messung erhobenen Daten analog dem Vorgehen bei Mausspots, wie die Bestimmung der Differenzmittelwerte, die Erstellung der Bland-Altman-Plots in relativen Zahlen sowie die Berechnung der 95%-Konfidenzgrenzen erwies sich aufgrund der aufgetretenen Unregelmäßigkeiten zwischen den Systemen KS Elispot und KS Elispot compact als nicht mehr sinnvoll und wurde daher nicht durchgeführt. Die von den verglichenen Lesesystemen R12 und R1 ermittelten relativen Anteile der Spots an den drei Durchmesserbereichen [20-50 $\mu$ m], [50-100 $\mu$ m] und [100-300 $\mu$ m] differierten so stark, dass ein Vergleich der beiden Reader anhand dieser Daten keine aussagekräftigen Ergebnisse erbracht hätte (Vgl. Tab. 24 und 3. 8. 1).

## 4. Diskussion

### 4.1. Allgemeine Aspekte

Die Elispot-Methode ist ein beliebtes Nachweisverfahren für spezifische T-Zell-Immunantworten. Seit ihrer Einführung im Jahr 1988 (Czerkinsky et al., 1988) wurde intensiv an der Optimierung, Standardisierung und Validierung der Elispot-Technik gearbeitet, so dass sie heute als eine der sensitivsten und zuverlässigsten Methoden zum Nachweis von Antigen-spezifischen T-Zellen auf Einzelzellebene gilt.

Im Gegensatz zum Elispot-Assay, an dessen Weiterentwicklung und Perfektionierung in den letzten Jahren kontinuierlich gearbeitet wurde, kam der Auswertung von Elispot-Proben bislang weit weniger Aufmerksamkeit zu. Die Elispot-Technik stellt positive T-Zell-Reaktionen als kleine Farbpunkte, so genannte Spots, mit einem Durchmesser von 20-300µm dar. Die Spots liegen in Vertiefungen, so genannten Wells, auf einer Elispot-Platte. Die Auswertung der Elispotplatten wurde lange Zeit durch Auszählen der Spots unter dem Stereomikroskop oder ohne Vergrößerungshilfe mit dem Auge vorgenommen. 1997 wurde von der Firma Zeiss das erste Bildanalysesystem zur automatisierten Auswertung von Elispot-Proben entwickelt.

In einer Vorarbeit konnte gezeigt werden, dass die Auswertung der Elispotplatten mit automatisierten Reader-Systemen der manuellen Auswertung unter dem Stereomikroskop oder mit dem Auge im Hinblick auf Standardisierbarkeit, Objektivierbarkeit und Reproduzierbarkeit deutlich überlegen ist (Janetzki et al., 2004). Die Arbeitsgruppe verglich als erste die unterschiedlichen Evaluationsmethoden von Elispot-Platten und untersuchte den Einfluss der Auswertungsmethode sowie auch den subjektiven Einfluss des Auswerter auf die Variabilität der Ergebnisse. Dazu wurden Elispot-Proben von drei Wissenschaftlern unter drei unterschiedlichen Evaluationsbedingungen ausgewertet, mit dem Stereomikroskop, mit einem automatisierten Readersystem, an dem die Untersucher die Spoterkennungsparameter selbst einstellten und mit einem automatisierten Readersystem, an dem die Parameter zur Spoterkennung bereits eingestellt waren und von den Untersuchern nicht verändert werden konnten. Die größte Variabilität der Ergebnisse zeigte sich in

der manuellen Auswertung von Elispot-Proben mit dem Stereomikroskop. Die hohe Variabilität war sowohl zwischen den einzelnen Untersuchern als auch innerhalb der Ergebnisse eines Auswerters festzustellen. Bei Auswertung der Elispot-Assays mit einem automatisierten Readersystem konnte eine deutliche Reduktion der Variabilität nachgewiesen werden. Damit belegt die Vorarbeit die Abhängigkeit der Ergebnisse des Elispot-Assays von der verwendeten Evaluationsmethode und zeigt die Notwendigkeit der Einbeziehung der Evaluationsmethode in die Standardisierungs- und Validierungsprozesse der einzelnen Arbeitsschritte der Elispot-Methode auf.

Inzwischen werden außer von der Firma Zeiss auch von verschiedenen anderen Firmen Geräte zur automatisierten Elispotprobenauswertung angeboten: A-El-Vis (Automated Elisa-Spot Assay Video Analysis Systems), AID (Autoimmun Diagnostika GmbH), Alpha Innotech Corp, Bio-Sys, CTL (Cellular Technology Ltd), Scanalytics Inc..

In dieser Arbeit wurden verschiedene Elispot-Reader-Systeme mit drei unterschiedlichen Auflösungsvermögen im Hinblick auf die Qualität ihrer Auswertung miteinander verglichen, das KS Elispot (R12) und das KS Elispot compact mit den Einstellungen 0.65-Zoom (R1) und 1.25-Zoom (R4). Damit wurde erstmals eine Vergleichsuntersuchung von verschiedenen automatisierten Bildanalyse-Systemen zur Auswertung von Elispotproben durchgeführt. Es existieren bislang keine Arbeiten, die bei der Diskussion der gewonnenen Ergebnisse herangezogen werden könnten. Die Resultate wurden daher solitär kommentiert.

#### **4.2. Aspekte zur Methodik**

Zur Beurteilung von Messgeräten werden die Begriffe Genauigkeit, Richtigkeit und Präzision verwendet, wobei Richtigkeit und Präzision verschiedene Aspekte der Genauigkeit darstellen. Die Begriffe wurden im Methodenteil genau definiert. Die Präzision gilt als ein Maß für die Übereinstimmung unabhängiger Messergebnisse unter festen Bedingungen. Sie gibt die Reproduzierbarkeit eines Messergebnisses an. In dieser Arbeit wurde die Präzision durch die Variabilität repräsentiert.

Die Richtigkeit und Genauigkeit wurden mittels der Bland-Alt-Plots visualisiert. Zur Abbildung der Richtigkeit, als ein Maß für die Abweichung zwischen dem Mittelwert einzelner Messergebnisse und dem wahren Wert der Messgröße, wurde der Mittelwert der Differenzen als Parallele zur x-Achse eingezeichnet. Die 95%-Konfidenzintervalle dienten zur Beurteilung der Genauigkeit, als Maß dafür, wie nahe das Ergebnis am wahren Wert der Messgröße lag.

### **4.3. Aspekte zu den Auswertungszeiten der Systeme**

Die in dieser Arbeit ermittelten Auswertungszeiten stimmten mit den vom Hersteller für beide Systeme bzw. für die drei Messanordnungen aufgeführten Zeitangaben überein.

Bei Verwendung nur eines Bildes pro Well (0.65-Zoom) wertete das KS Elispot compact 96 Wells einer Mikrotiterplatte in ca. 5 Minuten aus und damit über 6 Mal schneller als das KS Elispot. Wurden pro Well 4 Aufnahmen erstellt (1.25-Zoom), betrug die Auswertungszeit der 96 Wells am KS Elispot compact ca. 15 Minuten und damit weniger als die Hälfte der am KS Elispot benötigten Zeit.

Das KS Elispot benötigte mit 33 Minuten die mit Abstand längste Zeit zur Auswertung von 96 Wells. Demzufolge können während einer 7-stündigen Arbeitszeit am KS Elispot compact 0.65-Zoom maximal 84 Elispotplatten gelesen werden, am KS Elispot compact 1.25-Zoom können in der gleichen Zeit maximal 28 Elispot-Platten ausgewertet werden, während am KS Elispot maximal 13 Elispotplatten gemessen werden können. Hinzu kommt die Zeit, die zur Kalibrierung des Motortisches und zum (einmaligen) Erstellen des Settings sowie zum Auflegen bzw. Wechseln der Platten benötigt wird. Sie kann mit ca. 10 Minuten veranschlagt werden. Um die Effektivität der Systeme bezüglich des Faktors Zeit weiter zu steigern, werden Systeme angeboten, die über ein vollautomatisiertes Laden von Elispot-Platten (sog. „Plate-Loader“) verfügen. Sie ermöglichen damit einen hohen Plattendurchsatz pro Arbeitstag und entbinden den Untersucher von zeitintensiver Routinearbeit. Derartige Roboter sind allerdings auch deutlich teurer. Eine Voraussetzung für einen sinnvollen Einsatz besteht darin, dass alle Platten mit identischen Einstellungen gemessen werden können. Ansonsten wird die Einstellung der Parameter je Platte beliebig kompliziert.

Zusammenfassend weist das KS Elispot compact mit seiner 2-bzw. 6-fach schnelleren Auswertung einer kompletten 96-Well-Mikrotiterplatte gegenüber dem KS Elispot einen entscheidenden Vorteil auf. Insbesondere wenn große Mengen an Assay-Daten bearbeitet werden, resultiert eine immense Ersparnis an Zeit. Dies erfolgt allerdings auf Kosten der Messgenauigkeit (siehe auch 4. 5: Zuverlässigkeit).

#### **4.4. Aspekte zur Variabilität**

Die Werte für die Variabilität erstreckten sich über den Bereich von 2.3% bis 2.73%. Die niedrigste Variabilität mit einem Wert von 2.3% wurde für das System KS Elispot nachgewiesen, es folgte das System KS Elispot compact 1.25-Zoom mit einem Wert von 2.68%, dicht gefolgt vom System KS Elispot compact 0.65-Zoom, dessen Variabilität bei 2.73% lag. Bei beiden Elispot-Readern handelte es sich um automatisierte Lesesysteme, die hohe Reproduzierbarkeit der Ergebnisse konnte somit erwartet werden. Vom Hersteller wurde bezüglich der Variabilität für beide Systeme ein Wert unter 5% angegeben. Echte Unterschiede bezüglich der Variabilität und damit auch bezüglich der Präzision waren zwischen den Systemen nicht auszumachen.

#### **4.5. Aspekte zur Zuverlässigkeit der Spoterkennung**

Der Vergleich und die Auszählung der am KS Elispot aufgenommenen Well-Bilder ergaben, dass das KS Elispot 95.5% der echten Spots erkannt hatte. 4.5% der echten Spots wurden nicht erfasst. Eine wichtige Beobachtung war, dass in den 9 untersuchten Wells keine fälschlicherweise als positiv erkannten Spots nachzuweisen waren. Die Ursache dafür, dass einige Spots nicht erkannt wurden, lag in der überwiegenden Anzahl der Fälle darin, dass eng zusammen liegende Spots vom System als ein Spot quantifiziert worden waren. Nur in wenigen Fällen konnte das Nicht-Erkennen von Spots darauf zurückgeführt werden, dass ein isoliert gelegener, kleiner Spot nicht identifiziert worden war. Diese Ergebnisse lassen auf eine hohe Sensitivität der Spoterkennung durch das System KS Elispot schließen. Insbesondere die Tatsache, dass keine falsch-

positiven Spots gesehen wurden, spricht für eine zuverlässige Spoterkennung dieses Systems.

Das KS Elispot compact identifizierte bei Verwendung des 1.25-Zooms 84.29% der „wahren“ Spots, 15.71% wurden von diesem Reader-System nicht erkannt. Aus den Werten geht hervor, dass das KS Elispot compact weniger „wahre“ Spots als das KS Elispot erkannt hatte und sein Anteil an nicht erkannten Spots im Vergleich zum KS Elispot höher war. Der Anteil der falsch-positiven Spots an den Ergebnissen des KS Elispot compact 1.25-Zoom betrug 3.69%, während er für das System KS Elispot mit 0% angegeben werden konnte.

Mit der, nur ein Bild pro Well erstellenden, 0.65-Zoom-Einstellung erkannte das KS Elispot compact 79.36% der „wahren“ Spots und damit sowohl weniger als das KS Elispot als auch weniger als die 1.25-Zoom-Einstellung des KS Elispot compact. 20.64% der echten Spots wurden von diesem System nicht erfasst. Die falsch-positiven Spots machten einen Anteil von 5.85% am Gesamtergebnis aus. Dieser Wert lag über dem am KS Elispot compact 1.25-Zoom und am KS Elispot ermittelten Anteil an falsch-positiven Spots.

Verglichen mit dem KS Elispot erkannte das KS Elispot compact weniger „wahre“ Spots. Bei Verwendung des 0.65-Zooms und damit nur einer Aufnahme pro Well war der Anteil der erkannten „wahren“ Spots geringer, als bei Verwendung des 1.25-Zooms, bei dem 4 Bilder pro Well aufgenommen werden. Das KS Elispot compact zählte zum einen weniger Spots, da eng aneinander liegende Spots als ein großer identifiziert wurden, zum anderen, wesentlich häufiger als beim KS Elispot, wurden isoliert liegende Spot nicht erkannt. Es handelte sich dabei überwiegend um sehr kleine Spots des Durchmessers 20-35µm. Von entscheidender Bedeutung war weiterhin die Tatsache, dass das KS Elispot compact in weit mehr Fällen als das KS Elispot Schmutz oder andere Artefakte, bei denen es sich nicht um Spots handelte, als echte Spots identifiziert hatte. Dabei war der Anteil an falsch-positiven Spots am Gesamtergebnis bei Anfertigung von einer Aufnahme pro Well größer als bei Erstellung von vier Bildern.

Das KS Elispot compact blieb damit bezüglich der Zuverlässigkeit der Spoterkennung hinter dem KS Elispot zurück, für das die höchste Zuverlässigkeit bei der Spoterkennung nachgewiesen werden konnte.

## **4.6. Aspekte zu den Bland-Altman-Plots**

### **4.6.1. Rückschlüsse aus Differenzmittelwerten und Bland-Altman-Plots bei Maus-und Humanspots unabhängig von der Spotgröße**

Der für Mausspots bestimmte Mittelwert aller Differenzen wies beim Vergleich der Systeme KS Elispot (R12, 12 Bilder pro Well) gegen KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) als absoluter und relativer Wert in % ein positives Ergebnis auf. Veranschaulicht in den Bland-Altman-Plots lag der überwiegende Teil der Differenzwerte (R12-R4) oberhalb der Abszisse im positiven Wertebereich. Aus diesen Beobachtungen resultierte, dass die vom KS Elispot gemessenen Spotanzahlen überwiegend größer waren als die Werte, die vom KS Elispot compact 1.25-Zoom ermittelt wurden.

Beim Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) war der Mittelwert aller Differenzen als absolute Zahl positiv, während er als relative Zahl ein negatives Vorzeichen hatte. Dies konnte durch die folgende Tatsache erklärt werden: berechnete man die relativen Mittelwerte der Differenzen (R12-R1) für einzelne Spotanzahlbereiche, so war festzustellen, dass der Differenzmittelwert nur in der Kategorie [0-50] Spots ein negatives Ergebnis aufzeigte, während in allen anderen Spotanzahlbereichen positive Differenzmittelwerte vorlagen. Demnach hatten die Mittelwerte der Differenzen (R12-R1) bis auf den Spotanzahlbereich [0-50] positive Vorzeichen und es konnte folglich konstatiert werden, dass das KS Elispot mit Ausnahme des Spotbereichs [0-50] überwiegend mehr Spots als das KS Elispot compact 0.65-Zoom erkannt hatte. Bei Betrachtung des grafischen Verlaufes in den Bland-Altman-Plots wurde beim Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1) im Spotanzahlbereich [0-50] ein Großteil der Differenzen unterhalb der x-Achse gesehen. Dies war jedoch der einzige Bereich in dem

diese Verteilung nachgewiesen werden konnte, denn bei höheren Spotzahlen traten kaum mehr Werte unter 0 auf, der überwiegende Teil der Differenzen lag oberhalb der Abszisse im Positivbereich. Die vom KS Elispot gemessene Spotanzahl war demnach größer als die vom KS Elispot compact 0.65-Zoom gezählte.

Aus diesen Ergebnissen konnte zusammenfassend gefolgert werden, dass bei der Mausspotauswertung das System KS Elispot mehrheitlich mehr Spots erkannt hatte als das System KS Elispot compact. Eine Ausnahme war bei Wahl des 0.65-Zoom des KS Elispot compact (1 Bild pro Well) bei geringer Spotanzahl zu beobachten: bei Spotzahlen unter 50 pro Well zählte das KS Elispot etwa bei der Hälfte der Fälle weniger Spots als das KS Elispot compact. Sobald die Spotanzahl jedoch über 50 pro Well betrug, erkannte das KS Elispot fast immer mehr Spots als das KS Elispot compact.

Neben der Tatsache, dass das KS Elispot überwiegend mehr Spots als das KS Elispot compact gezählt hatte, war zu beobachten, dass ab einer Spotanzahl über 50 und insbesondere ab 100 Spots pro Well, die vom KS Elispot compact erkannte Spotmenge bei Verwendung nur eines Bildes pro Well deutlich stärker von der am KS Elispot gezählten Anzahl differierte als bei Verwendung von 4 Einzelbildern pro Well. Mit zunehmender Spotanzahl wurden vom KS Elispot compact unter Verwendung des 0.65-Zooms viel weniger Spots erkannt, als bei Wahl des 1.25-Zooms, bei dem die Differenzen zum System KS Elispot deutlich geringer ausfielen.

Die Mittelwerte der Differenzen (R12-R1) bzw. (R12-R4) dienen darüber hinaus der Abbildung der Richtigkeit, als Maß für die Abweichung zwischen dem Mittelwert einzelner Messgrößen und dem wahren Wert der Messgröße. Sie zeigten demnach an, wie gut der Mittelwert der Differenzen und der wahre Wert der Messgröße übereinstimmten. Der Mittelwert aller Differenzen (R12-R1) war höher als der Mittelwert aller Differenzen (R12-R4). Die größten Unterschiede zwischen den Differenzmittelwerten konnten bei Spotzahlen über 100 pro Well nachgewiesen werden, der Mittelwert der Differenzen (R12-R1) lag hier mit [33.57 bzw. 24.1%] klar über dem Mittelwert der Differenzen (R12-R4) mit [3.58 bzw. 2.6%]. Bezüglich der Richtigkeit erzielte das KS Elispot die besten

Ergebnisse, die Resultate des KS Elispot compact waren bei Verwendung von 4 Bildern pro Well deutlich besser als bei Aufnahme nur eines Bildes pro Well.

Die für die Humanspots bestimmten Mittelwerte der Differenzen (R12-R1) und (R12-R4) wiesen in ihren Angaben als absolute und relative Zahlen durchwegs positive Werte auf. Die Untersuchung der Verteilung der Differenzen (R12-R1) und (R12-R4) in den Bland-Altman-Graphiken zeigte, dass der Großteil der Differenzen oberhalb der x-Achse und demnach im positiven Wertebereich lag. Aus diesen Beobachtungen konnte gefolgert werden, dass die am KS Elispot (R12) gezählten Spotzahlen in den meisten Fällen größer als die am KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) und am KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) ermittelten Ergebnisse waren. Daraus konnte zunächst die allgemeine Schlussfolgerung gezogen werden, dass das KS Elispot über weite Strecken mehr Spots erkannt hatte als das KS Elispot compact.

Bei getrennter Betrachtung der beiden, am KS Elispot compact zur Verfügung stehenden Auflösungen 0.65-Zoom und 1.25-Zoom und deren Vergleich mit dem System KS Elispot, zeigten sich für beide unterschiedliche und von der Spotanzahl abhängige Ergebnisse:

bei einer Spotanzahl unter 400 pro Well differierten die Resultate des Vergleichs KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1), sowie die des Vergleichs KS Elispot gegen KS Elispot compact 1.25-Zoom (R4) kaum. Die jeweiligen Mittelwerte der Differenzen (R12-R1) bzw. (R12-R4) und graphischen Verläufe unterschieden sich nur geringfügig. Bei beiden Vergleichen erwies sich das KS Elispot als das System, das mehr Spots erkannt hatte, die Einstellungen 0.65-Zoom und 1.25-Zoom des KS Elispot compact erkannten dabei im ähnlichen Maß weniger Spots als das KS Elispot. Die Differenzmittelwerte betragen bei (R12-R1) [11.7 bzw. 26%], bei (R12-R4) [10.37 bzw. 21.4%].

Bei einer Spotanzahl von mehr als 400 pro Well, war es ebenfalls das KS Elispot, das in den beiden Vergleichen mehr Spots als das KS Elispot compact identifiziert hatte, es existierten jedoch deutliche Unterschiede zwischen der 0.65-Zoom-Einstellung des KS Elispot compact, bei der nur ein Bild pro Well erstellt wird und der 1.25-Zoom-Einstellung, bei der 4 Einzelaufnahmen pro Well angefertigt werden: die Abweichung in der vom KS Elispot und KS Elispot

compact erkannten Spotanzahl war bei Verwendung nur eines Bildes pro Well wesentlich größer als bei Erstellung von 4 Aufnahmen pro Well. Bei einer Aufnahme pro Well zählte das KS Elispot compact bei hohen Spotanzahlen merklich weniger Spots als bei Verwendung von 4 Einzelbildern pro Well. Die Differenzmittelwerte betragen für Spotzahlen über 400 bei (R12-R1) [104.5 bzw. 16.2%], bei (R12-R4) [29.3 bzw. 4.5%].

Die Mittelwerte der Differenzen dienen des Weiteren zur Visualisierung der Richtigkeit. Der Mittelwert aller Differenzen (R12-R1) lag über dem Mittelwert aller Differenzen (R12-R4). Bei Spotzahlen über 400 pro Well wurden die größten Unterschiede zwischen den Differenzmittelwerten nachgewiesen. Der Mittelwert der Differenzen (R12-R1) betrug [104.5 bzw. 16.2%], der Mittelwert der Differenzen (R12-R4) lag mit [29.3 bzw. 4.5%] deutlich darunter. Das KS Elispot compact erzielte grundsätzlich und insbesondere bei Spotzahlen über 400 pro Well bei Verwendung von 4 Bildern pro Well in Bezug auf die Richtigkeit bessere Ergebnisse als bei Verwendung eines Bildes pro Well.

Aus den Differenzmittelwerten und graphischen Verläufen in den Bland-Altman-Plots resultierte, dass das KS Elispot bei der Auswertung von Maus- und Human-Elispot-Proben in denselben Wells überwiegend mehr Spots identifiziert hatte als das KS Elispot compact. Weiterhin konnte beobachtet werden, dass bei steigenden Spotzahlen das KS Elispot compact bei Verwendung nur eines Bildes pro Well weniger Spots erkannt hatte als bei Erstellung von 4 Einzelbildern pro Well. Zudem erzielte das KS Elispot compact insbesondere bei hohen Spotzahlen pro Well bei Verwendung nur eines Bildes pro Well bezüglich der Richtigkeit schlechtere Ergebnisse als bei Erstellung von 4 Bildern pro Well. Zusammenfassend kann daher folgendes festgestellt werden:

- Das System KS Elispot erkannte in denselben Wells überwiegend mehr Spots als das System KS Elispot compact
- Das System KS Elispot compact identifizierte bei hohen Spotzahlen in den Wells bei Verwendung nur einer Aufnahme pro Well (0.65-Zoom-Einstellung) weniger Spots als bei 4 Aufnahmen pro Well (1.25-Zoom-Einstellung)

- Das System KS Elispot compact erzielte bei hohen Spotzahlen in Bezug auf die Richtigkeit bei Verwendung eines Bildes pro Well schlechtere Ergebnisse als bei 4 Aufnahmen pro Well
- Die deutlichsten Unterschiede zwischen den Systemen KS Elispot und KS Elispot compact wurden unter Verwendung der 0.65-Zoom-Einstellung des KS Elispot compact bei Spotzahlen über 400 pro Well bei Humanplatten und über 100 pro Well bei Mausplatten gesehen.

#### **4.6.2. Rückschlüsse aus Differenzmittelwerten und Bland-Altman-Plots bei Mausspots abhängig von der Spotgröße**

Die Auswertung der Mausspots an beiden Systemen unter Aufteilung der Spots in die zwei Spotdurchmesser-kategorien [20-100µm] und [100-300µm] erbrachte, dass die Mittelwerte der Differenzen (R12-R1) sowie (R12-R4) in absoluten und relativen Zahlen für beide Spotdurchmesserbereiche positiv waren. Demnach zählte das KS Elispot (R12, 12 Bilder pro Well) mehrheitlich mehr Spots als das KS Elispot compact 0.65-Zoom (R1, 1Bild pro Well) und 1.25-Zoom (R4, 4Bilder pro Well).

Die positiven Differenzmittelwerte unterschieden sich aber deutlich in ihren Werten: im Spotbereich [20-100µm] zeigte der Mittelwert aller Differenzen (R12-R1) [6.88 bzw. 24.73%] einen viel höheren Wert als der Mittelwert der Differenzen (R12-R4) [2.66 bzw. 14.26%]. Folglich zählte R1 bei Spots mit Durchmessern [20-100µm] weniger Spots als R4. Darüber hinaus fungierten die Mittelwerte der Differenzen zur Abbildung der Richtigkeit und zeigten demnach an, wie gut der Mittelwert der Differenzen und der wahre Wert der Messgröße übereinstimmten. Der Mittelwert der Differenzen (R12-R1) war deutlich größer als der Mittelwert der Differenzen (R12-R4). Daraus konnte abgeleitet werden, dass das KS Elispot compact bei Verwendung einer Aufnahme pro Well bezüglich der Richtigkeit schlechtere Resultate geliefert hatte als bei Erstellung von 4 Bildern pro Well.

Im Spotdurchmesserbereich [100-300µm] war der Mittelwert aller Differenzen (R12-R1) [0.48 bzw. 1.09%] kleiner als der Mittelwert aller Differenzen (R12-R4) [1.06 bzw. 9.4%]. Bei genauerer Betrachtung der einzelnen Differenzmittelwerte

(R12-R1), die für unterschiedliche Spotanzahlbereiche berechnet wurden, fiel jedoch auf, dass der Differenzmittelwert (R12-R1) ausschließlich im Spotbereich [0-50] kleiner war als der Differenzmittelwert (R12-R4). Schon im nächst höheren Spotanzahlbereich [50-100] lag der Mittelwert der Differenzen (R12-R1) [8.45 bzw. 12.11%] deutlich über dem Differenzmittelwert (R12-R4) [3.4 bzw. 5%]. Der niedrige Mittelwert aller Differenzen (R12-R1) war auf viele negative Werte im Spotanzahlbereich [0-50] zurückzuführen, die beim Vergleich R12 gegen R4 in diesem Bereich nicht in dem Ausmaß zu finden waren (siehe auch 3. 6. 5). Demnach zählte R1 bei Spots mit Durchmessern [100-300 $\mu$ m] weniger Spots als R4, ausgenommen bei Spotzahlen unter 50 pro Well. Weiterhin konnte aus den Mittelwerten der Differenzen, die zur Abbildung der Richtigkeit dienten, gefolgert werden, dass R1 in Bezug auf die Richtigkeit überwiegend schlechtere Ergebnisse erzielt hatte als R4.

Zusammenfassend sah man besonders bei den kleineren Spots [20-100 $\mu$ m] eine deutliche Divergenz in der vom KS Elispot und KS Elispot compact erkannten Spotanzahl; die am KS Elispot ermittelten Werte lagen dabei über den Ergebnissen des KS Elispot compact. Der Unterschied zwischen den beiden Systemen war hier größer als bei der alle Spotgrößen einbeziehenden Messung. Bei kleinen Spots zählte das KS Elispot compact bei Verwendung nur eines Bildes pro Well wesentlich weniger Spots als bei Verwendung von 4 Aufnahmen pro Well. Dieser Unterschied trat besonders bei Spotzahlen über 50 deutlich hervor. Wurden 4 Aufnahmen pro Well erstellt, so erkannte das KS Elispot compact zwar weniger Spots als das KS Elispot, doch lagen seine Ergebnisse deutlich näher an denen des KS Elispot, als die bei Verwendung nur eines Bildes pro Well ermittelten Ergebnisse. Der Einfluss der pro Well aufgenommenen Bilderzahl auf das Ergebnis des KS Elispot compact war bei kleinen Spots noch stärker zu beobachten als bei der spotgrößenunabhängigen Auswertung. In Bezug auf die Richtigkeit erzielte das KS Elispot bei Verwendung eines Bildes pro Well schlechtere Resultate als bei Erstellung von 4 Bildern pro Well. Die Unterschiede waren bei kleinen Spots deutlicher zu beobachten als bei der spotgrößenunabhängigen Messung.

Bei den größeren Spots [100-300µm] sah man zwar auch eine Abweichung in den Ergebnissen des KS Elispot und des KS Elispot compact, doch war diese wesentlich geringer als bei kleinen Spots und zudem auch geringer als bei der spotgrößenunabhängigen Auswertung der Mausspots. Die vom KS Elispot erkannte Spotanzahl lag dabei mehrheitlich über den Ergebnissen des KS Elispot compact.

Man konnte auch hier bei Spotzahlen über 50 aufzeigen, dass die vom KS Elispot compact gemessene Spotzahl bei Verwendung nur eines Bildes pro Well niedriger war als bei Verwendung von 4 Bildern pro Well. Der Unterschied zwischen den Ergebnissen des KS Elispot und des KS Elispot compact war demnach bei Erstellung von nur einer Aufnahme größer als bei 4 pro Well aufgenommen Bildern. Bezüglich der Richtigkeit lieferte das KS Elispot compact bei Erstellung nur einer Aufnahme pro Well schlechtere Ergebnisse als bei Verwendung von 4 Bildern pro Well, die Unterschiede waren bei größeren Spots jedoch geringer als bei Spots kleiner Durchmesser.

Zusammenfassend kann folgendes festgehalten werden:

- Das System KS Elispot erkannte bei kleinen Spotdurchmessern deutlich mehr Spots als das KS Elispot compact
- Bei kleinen Spotdurchmessern waren die Unterschiede zwischen KS Elispot und KS Elispot compact am größten, wenn am KS Elispot compact nur 1 Bild pro Well aufgenommen wurde (0.65-Zoom des KS Elispot compact)
- Die Unterschiede zwischen den Systemen waren bei kleinen Spots ausgeprägter als bei allen Spotgrößen
- Bei größeren Spotdurchmessern zählte das KS Elispot mehr Spots als das KS Elispot compact, die Systeme differierten jedoch wesentlich weniger in ihren Ergebnissen als bei kleinen Spotgrößen
- Bezüglich der Richtigkeit erzielte das KS Elispot compact bei Verwendung eines Bildes pro Well schlechtere Ergebnisse als bei Anfertigung von 4 Bildern pro Well. Bei kleinen Spots konnten größere Abweichungen in der Richtigkeit als bei großen nachgewiesen werden

#### **4.6.3. Rückschlüsse aus Differenzmittelwerten und Bland-Altman-Plots bei Humanspots in Abhängigkeit von der Spotgröße**

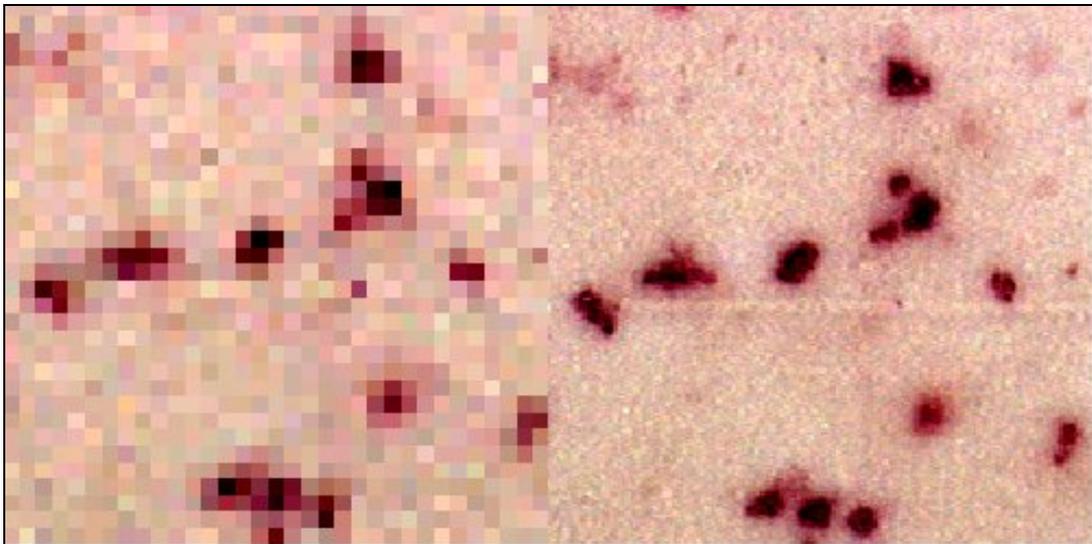
Im Vergleich zu den Mausplatten enthielten die Humanplatten mehr Spots sehr kleiner Durchmesser. Um die Unterschiede der Bildanalyzesysteme in Abhängigkeit von der Spotgröße zu untersuchen, wurden anstatt der bei Mausspots verwendeten zwei Durchmesserbereiche ([20-100µm], [100-300µm]) für Humanspots die drei Spotdurchmesser-kategorien, [20-50µm], [50-100µm] und [100-300µm] gewählt. Nachdem die aus diesen Daten angefertigten Bland-Altman-Plots für den Vergleich KS Elispot (R12) gegen KS Elispot compact 0.65-Zoom (R1) einen graphischen Verlauf zeigten, der von dem bis dahin beobachteten Trend völlig abwich, wurden die vom KS Elispot compact 0.65-Zoom ermittelten relativen Anteile der Spots in den drei Durchmesserbereichen mit den vom KS Elispot bestimmten Anteilen verglichen.

Dabei wurde eine gravierende Abweichung in den Werten festgestellt: im Durchmesserbereich [20-50µm] lag der vom KS Elispot compact ermittelte Spotanteil um das ca. 1.6-fache unter dem KS Elispot-Wert. Im Größenbereich [50-100µm] war der vom KS Elispot compact 0.65-Zoom ermittelte Spotanteil im Vergleich zum KS Elispot-Wert mehr als doppelt so hoch, im Durchmesserbereich [100-300µm] lag der vom KS Elispot compact 0.65-Zoom berechnete Spotanteil ca. 1.5-fach über dem KS Elispot-Wert. Das KS Elispot compact 0.65-Zoom hatte offenbar einen Großteil der kleinen Spots, die das KS Elispot noch dem Durchmesserbereich [20-50µm] zugeordnet hatte, schon im nächst höheren Durchmesserbereich [50-100µm] lokalisiert. Ebenso hatte das KS Elispot compact einen Teil der Spots, die das KS Elispot noch bei Durchmessern [50-100µm] lokalisierte, bereits dem nächst höheren Intervall [100-300µm] zugerechnet und damit die gesamte Verteilung der Größenanteile verschoben.

In den Bland-Altman-Diagrammen resultierte daraus für den Bereich [20-50µm] das Bild einer oberhalb der x-Achse linear ansteigenden Geraden. Die Differenzwerte (R12-R1) wurden mit zunehmender Spotanzahl immer größer. R1 erkannte viel weniger Spots als R12, da R1 diesem Durchmesserbereich fast 30% weniger Spots als R12 zugeordnet hatte. Für die Durchmesser [50-100µm] und [100-300] zeigten die Bland-Altman-Plots fast ausschließlich im

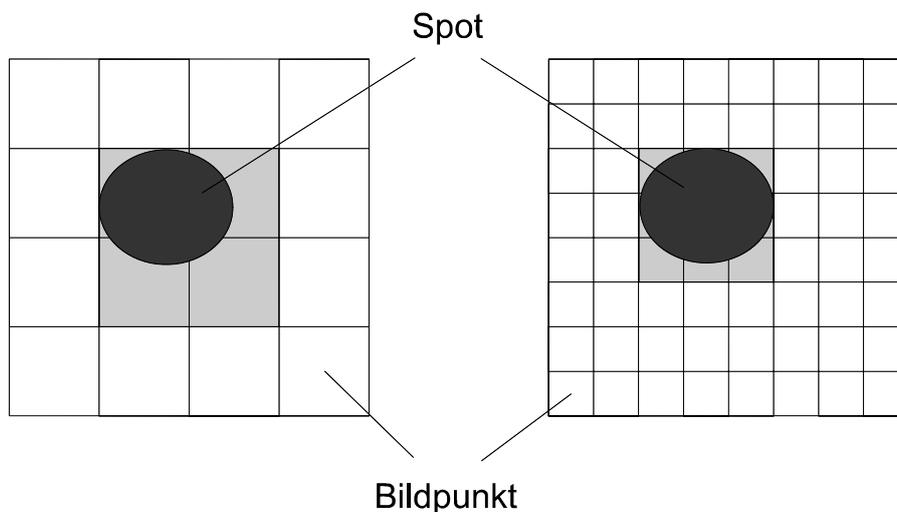
Negativbereich liegende Differenzen ( $R_{12}-R_1$ ), die mit zunehmender Spotzahl immer größer werdende negative Werte annahmen. Es ergab sich das Bild einer linear fallenden Geraden.  $R_{12}$  zählte in diesen Bereichen weniger Spots als  $R_1$ , da  $R_1$  diesen Durchmesserbereichen einen deutlich höheren Spotanteil als  $R_{12}$  zugeordnet hatte.

Als Ursache für dieses Phänomen konnten die unterschiedlichen Auflösungen der Systeme KS Elispot und KS Elispot compact angenommen werden (Abb. 45). Das KS Elispot nimmt mit dem Zeiss-Axioplan 2 imaging Mikroskop von jedem Well 12 Einzelbilder auf, während das KS Elispot compact bei Verwendung der 0.65-Zoom-Einstellung mit einem Stereomikroskop nur eine Aufnahme pro Well anfertigt. Für das KS Elispot compact beträgt die Pixelgröße bei einem  $760 \times 580$  Pixel Bild eines Wells ca.  $12 \mu\text{m}$ . Das KS Elispot erreicht über sein Auflichtmikroskop und die Aufnahme von 12 Einzelbildern pro Well eine deutlich höhere Auflösung, die Pixelgröße sinkt bei einem  $760 \times 580$  Pixel Bild eines Wells auf ca.  $2.6 \mu\text{m}$ .



**Abb. 45:** Die Aufnahme zeigt die unterschiedlichen Auflösungen der Systeme KS Elispot compact 0.65-Zoom (links, Pixelgröße ca.  $12 \mu\text{m}$ ) und KS Elispot (rechts, Pixelgröße ca.  $2.6 \mu\text{m}$ ) anhand von Bildausschnitten desselben Wells. Das KS Elispot compact 0.65-Zoom nimmt 1 Bild pro Well auf, das KS Elispot erstellt 12 Bilder pro Well.

Das System KS Elispot erreicht somit eine viel feinere Rasterung als das KS Elispot compact 0.65-Zoom. Die gröbere Rasterung des Compact-Systems hatte bei der Auswertung der kleinen Humanspots des Spotdurchmesserbereiches [20-50µm] dazu geführt, dass die Durchmesserangaben der Spots zum großen Teil über den tatsächlichen Werten lagen (Abb. 46). Dies führte zu der oben beschriebenen Zuordnung der Spots, die beim KS Elispot noch im kleineren Durchmesserbereich lagen, in den nächst höheren Größenbereich.



**Abb. 46:** Die vom Systemrechner interpretierte Spotgröße hängt von der Auflösung ab: bei niedriger Bildpunktauflösung (hohe Pixelgröße) liegt die ermittelte Größe kleiner Spots oft über dem tatsächlichen Wert, da der Systemrechner die Fläche auch nur teilweise vom Spot angeschnittener Bildpunkte addiert (linkes Bild). Bei höherer Bildpunktauflösung (geringe Pixelgröße) kann die Spotgröße exakter bestimmt werden (rechtes Bild).

## **4.7. Aspekte zu den 95%-Konfidenzgrenzen**

### **4.7.1. Aspekte zu den 95%-Konfidenzgrenzen für Mausspots**

Die 95%-Konfidenzintervalle geben den Bereich an, in dem 95% der Differenzen (R12-R1) bzw. (R12-R4) erwartet werden können. Darüber hinaus dienen sie

zur Beurteilung der Genauigkeit, als Maß dafür, wie nahe das Ergebnis am wahren Wert der Messgröße lag. Die Bestimmung der Konfidenzintervalle für einzelne Spotanzahlbereiche ermöglichte es, das entsprechende Intervall für eine bestimmte Spotanzahl heranzuziehen und so genauere Werte zu erhalten, als sie das für alle Spotanzahlen geltende Konfidenzintervall geliefert hätte. Interessiert z. B. der Vergleich von R12 gegen R1 bei einer Spotanzahl von ca. 100 pro Well, so kann gezielt am oberen und unteren Wert des für diesen Spotanzahlbereich geltenden Vertrauensintervalls abgelesen werden, wieviel Prozent R12 bei diesen Spotzahlen pro Well maximal mehr oder weniger zählte als R1.

Die aus den Mausdaten ermittelten 95%-Konfidenzgrenzen schlossen für den Vergleich KS Elispot (R12, 12 Bilder pro Well) mit KS Elispot compact 0.65-Zoom (R1, 1 Bild pro Well) das Intervall [-73.6%; 65.4%] ein. R12 erkannte demnach bis zu 65.4% mehr und bis zu 73.6% weniger Spots als R1. Dieses Ergebnis muss jedoch unter Berücksichtigung der für die einzelnen Spotanzahlbereiche bestimmten Vertrauensintervalle weiter erläutert werden:

bereits in vorangegangenen Teilen dieser Arbeit wurde aufgezeigt, dass beim Vergleich R12 gegen R1 in der durchmesserabhängigen und in der alle Spotgrößen einbeziehenden Messung bei Spotzahlen unter 50 viele negative Differenzen (R12-R1) aufgetreten waren. Die negativen Differenzen (R12-R1) waren überwiegend sehr klein, die Angabe der Differenzen als relative Zahlen in Prozent führte bei den geringen Spotzahlen jedoch oft zu hohen Werten, die sich dann auch in den für alle Spotzahlen geltenden 95%-Vertrauensintervallen niederschlugen. Es wurden daher zusätzlich die 95%-Vertrauensintervalle für die einzelnen Spotanzahlbereiche untersucht, um so auch Aussagen über die 95%-Konfidenzgrenzen bei Spotzahlen über 50 zu treffen. Dabei sah man, dass der untere Intervallwert (R12-R1) nur bei Spotzahlen unter 50 pro Well den deutlich negativen Wert [-80.6%] hatte. Im Spotbereich [50-100] lag der untere Intervallwert bei [-8.81%], im Spotbereich [100-200] betrug er [-3.33%]. Ausgenommen des Bereichs unter 50 Spots zählte das KS Elispot (R12) demnach maximal nur bis zu 8.81% (im Gegensatz zum unteren Intervallwert für alle Spotanzahlen [-73.6%]) weniger Spots pro Well als das KS Elispot 0.65-Zoom (R1). Die oberen Grenzen des 95%-Konfidenzintervalls (R12-R1) zeigten mit [50.35%] für Spotzahlen [0-50], mit [41.28%] für Spotzahlen [50-100] und mit

[51.62%] für Spotzahlen [100-200] durchwegs positive Werte. Das KS Elispot zählte folglich maximal bis zu 51.62% mehr Spots pro Well als das KS Elispot compact 0.65-Zoom. Die Einbeziehung der Vertrauensintervalle für die unterteilten Spotanzahlen macht deutlich, dass die isolierte Betrachtung des für alle Spotanzahlen geltenden Konfidenzintervalls zu einer ungenauen Interpretation führen kann und bestätigt deren zusätzlichen informativen Wert.

Für den Vergleich KS Elispot (R12) mit KS Elispot compact 1.25-Zoom (R4, 4 Bilder pro Well) lautete das 95%-Vertrauensintervall (R12-R4) [-32.5%; 52.9%]. R12 zählte demnach bis zu 52.9% mehr Spots und maximal 32.5% weniger Spots als R4. Für die drei Spotanzahlbereiche [0-50], [50-100] und [100-200] wurden für die untere Konfidenzgrenze (R12-R4) die Werte [-35.27%], [-17.61%] und [-19.9%] bestimmt. Die oberen Konfidenzgrenzen (R12-R4) waren mit [60.91%] für Spotzahlen [0-50], mit [29.9%] für Spotzahlen [50-100] und mit [25.23%] für Spotzahlen [100-200] durchwegs positiv. Bei Spotzahlen über 50 waren sie deutlich niedriger als die oberen Grenzen des (R12-R1)-Konfidenzintervalls mit [41.28%] für Spotzahlen [50-100] und mit [51.62%] für Spotzahlen [100-200]. R4 zählte demnach mehr Spots als R1.

Zusammenfassend konnte beobachtet werden, dass die oberen Grenzen der Vertrauensintervalle (R12-R1) und (R12-R4) deutlich über den unteren Grenzen lagen. Folglich war das Ausmaß, in dem das KS Elispot mehr Spots als das KS Elispot compact erkannte, deutlich größer als das, in dem es weniger als das Compact-System zählte. Für den Vergleich der Systeme bedeutete dies, dass das KS Elispot mehrheitlich mehr Spots erkannte hatte als das KS Elispot compact. Die oberen Grenzen der (R12-R1)-95%-Vertrauensintervalle wiesen höhere Werte auf als die der (R12-R4)-95%-Vertrauensintervalle. Das KS Elispot compact 0.65.Zoom (R1) maß folglich weniger Spots als das KS Elispot compact 1.25-Zoom (R4).

Die 95%-Konfidenzgrenzen (R12-R1) und (R12-R4) für kleine Mausspots mit Durchmessern [20-100 $\mu$ m] wiesen bezüglich ihrer oberen Intervallgrenzen positive Werte auf. Im Vergleich zu den oberen Grenzen der 95%-Vertrauensintervalle für alle Spotgrößen zeigten sie einen deutlichen Anstieg in ihren Werten, wobei die Zunahme beim (R12-R1)-Intervall stärker war. Demnach

war der Unterschied zwischen dem KS Elispot und dem KS Elispot compact bei kleinen Spots noch ausgeprägter nachzuweisen als bei allen Spotgrößen.

Bei großen Mausspots mit Durchmessern [100-300µm] waren die oberen Intervallgrenzen zwar auch positiv, verglichen mit den für alle Spotgrößen geltenden, waren sie jedoch niedriger. Für die unteren Intervallgrenzen konnten ebenso kleinere Werte nachgewiesen werden, so dass die Vertrauensintervalle insgesamt enger wurden. Demgemäß war der Unterschied zwischen dem KS Elispot und dem KS Elispot compact bei großen Spots wesentlich geringer als bei kleinen Spots und auch geringer als bei der alle Spotgrößen einbeziehenden Messung.

#### **4.7.2. Aspekte zu den 95%-Konfidenzgrenzen für Humanspots**

Das aus Humandaten ermittelte 95%-Konfidenzintervall (R12-R1) für den Vergleich KS Elispot (R12) mit KS Elispot compact 0.65-Zoom (R1) hatte die Intervallgrenzen [-19.2%; 71.4%]. Für den Vergleich KS Elispot (R12) mit KS Elispot compact 1.25-Zoom (R4) wurde das 95%-Konfidenzintervall (R12-R4) [-27.3%; 68.4%] bestimmt.

Das KS Elispot zählte demnach bis zu 71.4% mehr Spots und bis zu 19.2% weniger Spots als das KS Elispot compact 0.65-Zoom. Beim Vergleich mit dem KS Elispot compact 1.25 erkannte das KS Elispot bis zu 68.4% mehr Spots und maximal 27.3% weniger Spots als das Compact-System.

Die 95%-Vertrauensintervalle wurden auch für einzelne Spotanzahlbereiche bestimmt. Die oberen Grenzen der Vertrauensintervalle (R12-R1) und (R12-R4) wiesen dabei durchgehend positive Werte auf. Im Vergleich dazu zeigten die unteren Intervallgrenzen (R12-R1) und (R12-R4) ausnahmslos deutlich kleinere Werte, die größtenteils negativ waren. Es wurden aber auch positive untere Intervallgrenzen nachgewiesen. Bei positiver unterer und oberer Intervallgrenze maß R12 im jeweiligen Spotbereich immer mehr Spots als R1 bzw. R4.

Die Verteilung der Intervallgrenzen zeigte, dass das Ausmaß, in dem die KS Elispot-Ergebnisse über den Ergebnissen des Compact-Systems lagen, deutlich das Ausmaß übertraf, in dem sie darunter lagen. Demnach konnte die

Schlussfolgerung gezogen werden, dass das KS Elispot mehrheitlich mehr Spots als das KS Elispot compact erkannte.

Die größten Unterschiede zwischen den beiden Auflösungen des KS Elispot compact (0.65-bzw. 1.25-Zoom) wurden bei Humanspots bei Spotzahlen über 400 gesehen: das KS Elispot hatte beim Vergleich mit dem KS Elispot compact 0.65-Zoom, das nur 1 Bild pro Well erstellt, im Spotbereich [400-600] bis zu 41.77% und im Spotbereich [600-900] bis zu 30.51% mehr Spots gezählt, während es beim Vergleich mit KS Elispot compact 1.25-Zoom, das 4 Bilder pro Well aufnimmt, im Spotbereich [400-600] nur bis zu 27.3% und im Spotbereich [600-900] nur bis zu 10.64% mehr Spots identifizierte.

Aus den Ergebnissen der 95%-Konfidenzintervalle für Maus-und Humanspots resultierten die folgenden Aspekte:

- Beim Vergleich KS Elispot gegen KS Elispot compact wiesen die oberen Grenzen der 95%-Vertrauensintervalle durchwegs positive und deutlich höhere Werte auf als die unteren Intervallgrenzen, die größtenteils negative Werte hatten. Das Ausmaß, in dem KS Elispot mehr Spots als KS Elispot compact erkannt hatte, war deutlich größer als das, in dem es weniger Spots als KS Elispot compact gezählt hatte.
- Die oberen Intervallgrenzen für den Vergleich KS Elispot gegen KS Elispot compact lagen bei Verwendung eines Bildes pro Well höher als bei Erstellung von 4 Aufnahmen pro Well. Das KS Elispot compact zählte bei Verwendung des 0.65-Zooms weniger Spots als bei Verwendung des 1.25-Zooms.
- Bei kleinen Spotdurchmessern [20-100µm] waren die 95%-Vertrauensintervalle weiter und zeigten höhere Werte für die oberen Intervallgrenzen als die für alle Spotgrößen und für große Spots ermittelten 95%-Vertrauensintervalle. Die Unterschiede zwischen den Systemen waren bei kleinen Spots größer als bei großen Spots.
- Bei großen Spotdurchmessern [100-300µm] waren die 95%-Vertrauensintervalle enger und die oberen Grenzen der 95%-Vertrauensintervalle kleiner als die der für alle Spotgrößen und für kleine Spots ermittelten 95%-Vertrauensintervalle. Die Unterschiede zwischen den Systemen waren bei großen Spots geringer als bei kleinen Spots.

Die über die 95%-Konfidenzgrenzen getroffenen Aussagen stehen im Einklang mit den Schlussfolgerungen, die aus dem Vergleich der Differenzmittelwerte und aus dem graphischen Verlauf in den Bland-Altman-Plots gezogen wurden. Die Gründe für die beschriebenen Resultate und für die Unterschiede zwischen den beiden Systemen werden im Folgenden erörtert (4. 8).

#### **4.8. Aspekte zu Unterschieden zwischen den Systemen**

Um den Vergleich der Systeme KS Elispot und KS Elispot compact visualisieren und quantifizieren zu können, wurden mehrere Methoden eingesetzt. Die aus ihnen gewonnenen Ergebnisse und Schlussfolgerungen stimmten in ihren Aussagen überein.

Demnach erkannte das KS Elispot über weite Strecken mehr Spots als das KS Elispot compact, wobei das KS Elispot compact bei Verwendung nur eines Bildes pro Well (0.65-Zoom) schlechtere Ergebnisse erzielte als bei 4 Aufnahmen pro Well (1.25-Zoom).

Des Weiteren erwies sich das KS Elispot im Vergleich zum KS Elispot compact als zuverlässiger bei der Spoterkennung. Auch hier konnte nachgewiesen werden, dass die Auswertung des KS Elispot compact bei Aufnahme von 4 Bildern pro Well zuverlässiger war als bei Erstellung nur eines Bildes pro Well: bezüglich des Anteils an falsch-positiven Spots am Gesamtergebnis hatte das KS Elispot compact bei Verwendung einer Aufnahme pro Well den höchsten Wert der Systeme bzw. der drei Messanordnungen, sein Anteil an richtig erkannten echten Spots war am niedrigsten. Bei Aufnahme von 4 Bildern pro Well konnten bessere Ergebnisse erzielt werden: der Anteil an falsch-positiven Spots am Ergebnis war kleiner und der Anteil der richtig erkannten echten Spots stieg. Die Werte lagen jedoch immer noch unterhalb der Ergebnisse, die das KS Elispot erzielt hatte, das von allen Systemen den wahren Werten am nächsten kam.

Als ursächlicher Faktor für diese Ergebnisse ist die ungleiche Auflösung der 2 Systeme bzw. der drei Messanordnungen anzusehen. Das KS Elispot ist mit einem Auflichtmikroskop, dem Zeiss Axioplan 2 imaging mit Epiplan Neofluar 5x / 0,15 Objektiv und Zoomadapter ausgestattet. Über dieses Mikroskop werden

pro Well 12 Einzelbilder aufgenommen. Aufgrund der so erreichten hohen Auflösung beträgt die Pixelgröße bei einem 760x580 Pixel Bild eines Wells ca. 2.6µm.

Das KS Elispot compact verwendet ein Stereomikroskop, das Zeiss Stemi 2000 C mit Auflichtbeleuchtung. Bei Wahl des 0.65-Zooms wird von jedem Well ein Bild aufgenommen. Es resultiert eine Pixelauflösung von 12µm. Die Auflösung kann durch Wahl des 1.25-Zooms gesteigert werden. Bei dieser Einstellung werden pro Well 4 Einzelaufnahmen angefertigt. Die Systemauflösung verdoppelt sich, es wird eine Pixelgröße von 6µm erreicht.

Die Gegenüberstellung der Auflösungen des KS Elispot und des KS Elispot compact 0.65-Zoom anhand einer an beiden Systemen erstellten Aufnahme desselben Wells (Abb. 45, Seite 115) veranschaulicht sehr klar, dass eine geringe Auflösung Probleme in folgenden Bereichen nach sich ziehen kann:

- Erkennung kleiner Spots
- Unterscheidung echter Spots von Artefakten und Schmutz anhand der typischen Spotmerkmale
- Sichere Unterscheidung eng zusammen liegender Spots
- Exakte Bestimmung der Spotgröße

Die in dieser Arbeit aus dem Vergleich der Systeme gewonnenen Resultate können diese Annahme bestätigen. Sie zeigen, dass das KS Elispot compact 0.65-Zoom, im Vergleich zum KS Elispot compact 1.25-Zoom und zum KS Elispot, insbesondere bei kleinen Spots schlechtere Messergebnisse lieferte. Bei großen Spots lag das KS Elispot compact 0.65-Zoom bezüglich der erkannten Spotanzahl zwar auch hinter den zwei anderen Systemen zurück, doch nicht in dem Maße wie bei Spots kleiner Durchmesser.

Neben der Größe der Spots war die Spotanzahl für das Messergebnis bedeutsam. Bei hohen Spotzahlen pro Well (Human >400, Maus >100) war die Auswertung des KS Elispot compact 0.65-Zoom schlechter als die der anderen Systeme. Bei hohen Spotzahlen sind die einzelnen Spots oft nicht gleichmäßig und klar voneinander getrennt auf dem Nitrozelluloseboden verteilt, sondern können aneinander grenzen und überlappen. Ein System mit geringerer

Auflösung kann eng zusammenliegende Spots oft nicht klar identifizieren und erfasst sie als einen großen Spot.

Zusätzlich zur Auflösung kann als weiterer Faktor die Sorgfalt, mit der die einzelnen Arbeitsschritte der Elispot-Methode durchgeführt wurden, die Ergebnisse der Systeme beeinflussen. Die resultierende Qualität der Spots auf den Böden der Mikrotiterplatten ist für eine optimale Spoterkennung durch das Auswertungssystem von großer Wichtigkeit. Im idealen Fall wurden die Elispot-Platten so präpariert, dass die Spots die typische Merkmale aufweisen: sie sind rund, gehen am Rand aber fließend in den Hintergrund über und haben somit keine scharfe Begrenzung, ihre Farbtintensität ist im Zentrum am höchsten und nimmt zum Rand hin kontinuierlich ab. Um starke Überlagerungen zu vermeiden, sollten die Spots zudem gleichmäßig und in einer nicht zu hohen Anzahl über das Well verteilt sein. Weiterhin sollten die Platten für ein optimales Lesen durch die automatisierten Systeme kaum Artefakte, Schmutz und Untergrundfärbung aufweisen.

Die Nitrozelluloseböden der in dieser Arbeit verwendeten Humanplatten wiesen Spots mit sehr gut ausgeprägten, typischen Spotmerkmalen auf, sie zeigten relativ wenige Artefakte und hatten keine Untergrundfärbung. Bei den Mausplatten war eine leichte Untergrundfärbung festzustellen, die Verteilung der Spots war etwas unregelmäßig und die Anzahl der Artefakte lag höher als bei den Humanplatten.

Aus den Datenauswertungen war zu entnehmen, dass die größten Unterschiede zwischen den beiden Zoom-Einstellungen 0.65 und 1.25 des KS Elispot compact bei Mausplatten schon bei Spotzahlen über 100 gesehen wurden, während sie bei Humanplatten erst bei Spotzahlen über 400 auftraten. Die Unterschiede zwischen den Systemen waren bei Mausspots bereits bei kleineren Spotanzahlen pro Well erkennbar, obwohl die Maus-Elispot-Proben im Vergleich zu den Humanplatten geringere Spotanzahlen pro Well aufzeigten und der Anteil kleiner Spotdurchmesser bei ihnen niedriger als bei den Elispot-Proben mit Humanspots lag. Dies konnte vorrangig dadurch erklärt werden, dass die Mausplatten mehr Artefakte, Schmutz und eine stärkere Untergrundfärbung als die Humanplatten aufwiesen. Zudem kann es durch die größeren Durchmesser der Mausspots bei der gleichzeitig unregelmäßigen Spotverteilung leichter als

bei Human-Proben zu Überlappungen gekommen sein. Die Anzahl der Humanspots pro Well war zwar höher, doch hatten diese kleinere Durchmesser und waren regelmäßig auf den Nitrocelluloseböden verteilt. Demzufolge hängt eine gute Auswertung von Elispot-Platten durch automatisierte Bildanalysesysteme nicht zuletzt auch von einer optimalen Präparation der verwendeten Elispot-Proben ab.

Diese Arbeit belegt den Einfluss des verwendeten automatisierten Auswertungssystems auf das Ergebnis der Elispot-Methode. Sie zeigt weiterhin die Abhängigkeit einer optimalen Auswertung von der Präparation der Elispot-Proben auf. Dies sollte in der Standardisierung und Validierung der Methode mitberücksichtigt werden.

## 5. Zusammenfassung

Die vorliegende Arbeit hatte das Ziel, erstmals unterschiedliche automatisierte Bildanalyzesysteme in der Auswertung von Elispot-Proben miteinander zu vergleichen. Neben dem Vergleich der Anzahl der gemessenen Spots, sollten die Zeit für die Messungen, die Genauigkeit, Richtigkeit und Präzision untersucht werden. Weiterhin sollten die Zuverlässigkeit der Messungen sowie die Einflussgrößen der einzelnen Bildanalyzesysteme geprüft werden. Es wurden Elispotproben von verschiedenen Spezies (Maus und Human) und von verschiedenen unabhängigen Arbeitsgruppen verwendet.

Die verglichenen Bildanalyzesysteme bestanden aus den aufeinander abgestimmten Komponenten Mikroskop, Farbkamera, Motortisch mit Steuerung, Analysesoftware und Rechner. Sie unterschieden sich in ihren Mikroskopen, in der Anzahl der pro Well aufgenommenen Bilder und in der damit erreichten Bildpunktauflösung. Beim KS Elispot wurden im Auflichtmikroskop pro Well 12 Bilder über eine Farbkamera aufgenommen, die Bildpunktauflösung für ein 760x580 Pixel Bild eines Wells betrug 2.6µm. Beim KS Elispot compact 0.65-Zoom-Einstellung wurde im Stereomikroskop pro Well 1 Bild über eine Farbkamera aufgenommen und eine Bildpunktauflösung von 12µm erreicht. Beim KS Elispot 1.25-Zoom-Einstellung wurden im Stereomikroskop 4 Bilder pro Well über eine Farbkamera erstellt, die Bildpunktauflösung betrug 6µm.

Im Bezug auf den Faktor Zeit war das KS Elispot compact dem KS Elispot deutlich überlegen. Bei Verwendung eines Bildes pro Well bzw. 4 Bildern pro Well wertete das KS Elispot compact eine komplette 96-Well-Mikrotiterplatte bis zu 6 Mal schneller bzw. mindestens doppelt so schnell aus wie das KS Elispot. Die hohe Auflösung des KS Elispot resultiert in einer langen Auswertungszeit der einzelnen Platten. Werden große Mengen von Elispot-Proben ausgewertet, so bietet das KS Elispot compact aufgrund seiner kürzeren Messzeiten eine deutliche Ersparnis an Zeit und damit gegenüber dem KS Elispot unter diesem Gesichtspunkt einen entscheidenden Vorteil. Die Variabilität der Messungen lag bei allen Systemen niedrig, ohne nennenswerte Unterschiede zwischen den Systemen. Die höchste Zuverlässigkeit bei der Spoterkennung konnte für das System KS Elispot nachgewiesen werden. Es erkannte nahezu alle echten Spots und mehr echte Spots als das KS Elispot compact. Das KS Elispot wies im

Gegensatz zum KS Elispot compact keine falsch-positiven Spots auf. Bei Verwendung von 4 Bildern pro Well arbeitete das KS Elispot compact zuverlässiger als bei Erstellung von nur einer Aufnahme pro Well: der Anteil an falsch-positiven Spots am Ergebnis sank und der Anteil der richtig erkannten Spots stieg. Die Werte lagen jedoch immer noch unterhalb der Ergebnisse, die das KS Elispot erzielt hatte.

Das KS Elispot compact erkannte grundsätzlich in denselben Wells weniger Spots als das KS Elispot, das der tatsächlichen Anzahl der Spots am nächsten kam. Bei Verwendung von nur einer Aufnahme pro Well identifizierte das KS Elispot compact deutlich weniger Spots als bei Aufnahme von 4 Bildern pro Well. Die deutlichsten Unterschiede zwischen den beiden Systemen KS Elispot und KS Elispot compact wurden bei der Messung durch das KS Elispot compact mit einem Bild pro Well bei Spotzahlen über 100 bei Mausspots und über 400 bei Humanspots gesehen. Die zwischen dem KS Elispot und dem KS Elispot compact nachgewiesenen Unterschiede waren bei kleinen Spots (bis 100µm) deutlich größer als bei Spots größerer Durchmesser.

Der Vergleich der Systeme erbrachte, dass das hochauflösende KS Elispot eine bessere Auswertungsqualität als das KS Elispot compact bietet, insbesondere bei der Auswertung von Elispot-Proben, die sehr viele und zudem sehr kleine Spots enthielten. Beim KS Elispot compact war die Messung mit 4 Bildern pro Well der Auswertung mit nur einem Bild pro Well bezüglich der Zuverlässigkeit klar überlegen. Bei Verwendung des KS Elispot compact sollte deshalb bei sehr kleinen Spots zumindest die mit 4 Bildern pro Well arbeitende Einstellung gewählt werden. Weiterhin ist zu bemerken, dass eine optimale Auswertung von Elispot-Proben durch automatisierte Reader-Systeme maßgeblich durch die Präparation der verwendeten Elispot-Proben und die daraus resultierende Qualität der Spots beeinflusst wird. Zahlreiche Artefakte, eine starke Untergrundfärbung oder nicht deutlich ausgebildete typische Spotmerkmale können die Messergebnisse eines Systems beeinträchtigen. Hierbei wurde das KS Elispot compact System stärker beeinflusst als das KS Elispot System.

In dieser Arbeit konnte nachgewiesen werden, dass die Auswertung der Elispot-Proben und damit die gewonnenen Ergebnisse von dem verwendeten, automatisierten Lesesystem abhängen. Die hier dargelegten Informationen über das Auswertungsvermögen der verschiedenen Reader-Systeme können

individuell genutzt werden. Die Arbeit unterstreicht den hohen Stellenwert der Standardisierung, Validierung und Optimierung aller Komponenten der Elispot-Methode. Dies ist auch gerade in Bezug auf die Weiterentwicklung dieser Technik und die Eröffnung von weiteren Einsatzspektren unerlässlich.

## 6. Literatur

- Abbas AK, Lichtman AH. 2004. Basic Immunology. Elsevier.
- Andersen MH, Pedersen LO, Becker JC, Straten PT. 2001. Identification of a cytotoxic T lymphocyte response to the apoptosis inhibitor protein survivin in cancer patients. *Cancer Res* 61:869-872.
- Anthony DD, Post AB, Valdez H, Peterson DL, Murphy M, Heeger PS. 2001. ELISPOT analysis of hepatitis C virus protein-specific IFN-gamma-producing peripheral blood lymphocytes in infected humans with and without cirrhosis. *Clin Immunol* 99:232-240.
- Bennouna J, Hildesheim A, Chikamatsu K, Gooding W, Storkus WJ, Whiteside TL. 2002. Application of IL-5 ELISPOT assays to quantification of antigen-specific T helper responses. *J Immunol Methods* 261:145-156.
- Berger T, Haendle I, Schrama D, Luftl M, Bauer N, Pedersen L, Schuler-Thurner B, Hohenberger W, Straten PP, Schuler G, Becker J. 2004. Circulation and homing of melanoma-reactive T cells to both cutaneous and visceral metastases after vaccination with monocyte-derived dendritic cells. *Int J Cancer* 111:229-237.
- Bland J, Altman D. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307-310.
- Bland J, Altman D. 1995. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 346:1085-1087.
- Bland J, Altman D. 1999. Measuring agreement in method comparison studies. *Stat Methods Med Res* 8:135-160.
- Bland J, Altman D. 2003. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol* 22:85-93.
- Bontems P, Robert F, Van Gossum A, Cadranel S, Mascart F. 2003. Helicobacter pylori modulation of gastric and duodenal mucosal T cell cytokine secretions in children compared with adults. *Helicobacter* 8:216-226.
- Britten CM, Meyer RG, Kreer T, Drexler I, Wolfel T, Herr W. 2002. The use of HLA-A\*0201-transfected K562 as standard antigen-presenting cells for CD8(+) T lymphocytes in IFN-gamma ELISPOT assays. *J Immunol Methods* 259:95-110.
- Carvalho LH, Hafalla JC, Zavala F. 2001. ELISPOT assay to measure antigen-specific murine CD8(+) T cell responses. *J Immunol Methods* 252:207-218.
- Clay T, Hobeika A, Mosca P, Lyerly H, Morse M. 2001. Assays for monitoring cellular immune responses to active immunotherapy of cancer. *Clin Cancer Res* 7:1127-1135.

- Corne P, Huguet MF, Briant L, Segondy M, Reynes J, Vendrell JP. 1999. Detection and enumeration of HIV-1-producing cells by ELISPOT (enzyme-linked immunospot) assay. *J Acquir Immune Defic Syndr Hum Retrovirol* 20:442-447.
- Coulie P. 1997. Human tumour antigens recognized by T cells: new perspectives for anti-cancer vaccines? *Mol Med Today* 3:261-268.
- Cui Y, Chang L. 1997. Computer-assisted, quantitative cytokine enzyme-linked immunospot analysis of human immune effector cell function. *Biotechniques* 22:1146-1149.
- Czerkinsky C, Andersson G, Ekre HP, Nilsson LA, Klareskog L, Ouchterlony O. 1988. Reverse ELISPOT assay for clonal analysis of cytokine production. I. Enumeration of gamma-interferon-secreting cells. *J Immunol Methods* 110:29-36.
- Czerkinsky CC, Nilsson LA, Nygren H, Ouchterlony O, Tarkowski A. 1983. A solid-phase enzyme-linked immunospot (ELISPOT) assay for enumeration of specific antibody-secreting cells. *J Immunol Methods* 65:109-121.
- Davidson EJ, Sehr P, Faulkner RL, Parish JL, Gaston K, Moore RA, Pawlita M, Kitchener HC, Stern PL. 2003. Human papillomavirus type 16 E2- and L1-specific serological and T-cell responses in women with vulval intraepithelial neoplasia. *J Gen Virol* 84:2089-2097.
- Deutsches Institut für Normung (DIN) B, Berlin. 1997. DIN 55320, Teil 17, (1997), Begriffe zur Qualitätssicherung und Statistik. Begriffe zur Genauigkeit von Ermittlungsverfahren und Ermittlungsergebnissen.
- Dewitte K, Fierens C, Stockl D, Thienpont L. 2002. Application of the Bland-Altman plot for interpretation of method-comparison studies: a critical investigation of its practice. *Clin Chem* 48:799-801; author reply 801-792.
- Eming R, Budinger L, Riechers R, Christensen O, Bohlen H, Kalish R, Hertl M. 2000. Frequency analysis of autoreactive T-helper 1 and 2 cells in bullous pemphigoid and pemphigus vulgaris by enzyme-linked immunospot assay. *Br J Dermatol* 143:1279-1282.
- Ewer K, Deeks J, Alvarez L, Bryant G, Waller S, Andersen P, Monk P, Lalvani A. 2003. Comparison of T-cell-based assay with tuberculin skin test for diagnosis of *Mycobacterium tuberculosis* infection in a school tuberculosis outbreak. *Lancet* 361:1168-1173.
- Farina C, Then Bergh F, Albrecht H, Meinl E, Yassouridis A, Neuhaus O, Hohlfeld R. 2001. Treatment of multiple sclerosis with Copaxone (COP): Elispot assay detects COP-induced interleukin-4 and interferon-gamma response in blood cells. *Brain* 124:705-719.
- Fujihashi K, McGhee JR, Beagley KW, McPherson DT, McPherson SA, Huang CM, Kiyono H. 1993. Cytokine-specific ELISPOT assay. Single cell analysis of IL-2, IL-4 and IL-6 producing cells. *J Immunol Methods* 160:181-189.

- Gabrielsson S, Soderlund A, Paulie S, van der Pouw Kraan TC, Troye-Blomberg M, Rak S. 2001. Specific immunotherapy prevents increased levels of allergen-specific IL-4- and IL-13-producing cells during pollen season. *Allergy* 56:293-300.
- Gajewski TF. 2000. Monitoring specific T-cell responses to melanoma vaccines: ELISPOT, tetramers, and beyond. *Clin Diagn Lab Immunol* 7:141-144.
- Godard B, Gazagne A, Gey A, Baptiste M, Vingert B, Pegaz-Fiornet B, Strompf L, Fridman W, Glotz D, Tartour E. 2004. Optimization of an elispot assay to detect cytomegalovirus-specific CD8+ T lymphocytes [In Process Citation]. *Hum Immunol* 65:1307-1318
- Griffioen M, Borghi M, Schrier PI, Osanto S. 2001. Detection and quantification of CD8(+) T cells specific for HLA-A\*0201-binding melanoma and viral peptides by the IFN-gamma-ELISPOT assay. *Int J Cancer* 93:549-555.
- Harms V. 1997. *Biomathematik, Statistik und Dokumentation*. Harms Verlag Kiel, 1997.
- Hellings N, Baree M, Verhoeven C, D'Hooghe M B, Medaer R, Bernard CC, Raus J, Stinissen P. 2001. T-cell reactivity to multiple myelin antigens in multiple sclerosis patients and healthy controls. *J Neurosci Res* 63:290-302.
- Herr W, Linn B, Leister N, Wandel E, Meyer zum Buschenfelde KH, Wolfel T. 1997. The use of computer-assisted video image analysis for the quantification of CD8+ T lymphocytes producing tumor necrosis factor alpha spots in response to peptide antigens. *J Immunol Methods* 203:141-152.
- Herr W, Schneider J, Lohse AW, Meyer zum Buschenfelde KH, Wolfel T. 1996. Detection and quantification of blood-derived CD8+ T lymphocytes secreting tumor necrosis factor alpha in response to HLA-A2.1-binding melanoma and viral peptide antigens. *J Immunol Methods* 191:131-142.
- Hill P, Fox A, Jeffries D, Jackson-Sillah D, Lugos M, Owiafe P, Donkor S, Hammond A, Corrah T, Adegbola R, McAdam K, Brookes R. 2005. Quantitative T cell assay reflects infectious load of Mycobacterium tuberculosis in an endemic case contact model [In Process Citation]. *Clin Infect Dis* 40:273-278
- Hricik DE, Rodriguez V, Riley J, Bryan K, Tary-Lehmann M, Greenspan N, DeJelo C, Schulak JA, Heeger PS. 2003. Enzyme linked immunosorbent spot (ELISPOT) assay for interferon-gamma independently predicts renal function in kidney transplant recipients. *Am J Transplant* 3:878-884.
- Janetzki S, Schaed S, Blachere N, Ben-Porat L, Houghton A, Panageas K. 2004. Evaluation of Elispot assays: influence of method and operator on variability of results. *J Immunol Methods* 291:175-183.
- Janeway CA, Travers P, Walport M. 2002. *Immunologie*. Spektrum Akademischer Verlag, 2002.

- Kabilan L, Andersson G, Lolli F, Ekre HP, Olsson T, Troye-Blomberg M. 1990. Detection of intracellular expression and secretion of interferon-gamma at the single-cell level after activation of human T cells with tetanus toxoid in vitro. *Eur J Immunol* 20:1085-1089.
- Khan N, Cobbold M, Keenan R, Moss PA. 2002. Comparative analysis of CD8+ T cell responses against human cytomegalovirus proteins pp65 and immediate early 1 shows similarities in precursor frequency, oligoclonality, and phenotype. *J Infect Dis* 185:1025-1034.
- Kiessling A, Schmitz M, Stevanovic S, Weigle B, Holig K, Fussel M, Fussel S, Meye A, Wirth MP, Rieber EP. 2002. Prostate stem cell antigen: Identification of immunogenic peptides and assessment of reactive CD8+ T cells in prostate cancer patients. *Int J Cancer* 102:390-397.
- Koch KR. 1987. Parameterschätzung und Hypothesentest. Ferd. Dümmlers Verlag, Bonn, 1987.
- Kreyszig E. 1968. Statistische Methoden und ihre Anwendung. Vadenhoek und Ruprecht Verlag, Göttingen, 1968.
- Krouwer J, Monti K. 1995. A simple, graphical method to evaluate laboratory assays. *Eur J Clin Chem Clin Biochem* 33:525-527.
- Kumar A, Weiss W, Tine JA, Hoffman SL, Rogers WO. 2001. ELISPOT assay for detection of peptide specific interferon-gamma secreting cells in rhesus macaques. *J Immunol Methods* 247:49-60.
- Lalvani A, Pathan AA, Durkan H, Wilkinson KA, Whelan A, Deeks JJ, Reece WH, Latif M, Pasvol G, Hill AV. 2001. Enhanced contact tracing and spatial tracking of Mycobacterium tuberculosis infection by enumeration of antigen-specific T cells. *Lancet* 357:2017-2021.
- Lee FK, Nahmias AJ, Lowery S, Nesheim S, Reef S, Thompson S, Oleske J, Vahlne A, Czerkinsky C. 1989. ELISPOT: a new approach to studying the dynamics of virus-immune system interaction for diagnosis and monitoring of HIV infection. *AIDS Res Hum Retroviruses* 5:517-523.
- Lewis JJ, Janetzki S, Schaed S, Panageas KS, Wang S, Williams L, Meyers M, Butterworth L, Livingston PO, Chapman PB, Houghton AN. 2000. Evaluation of CD8(+) T-cell frequencies by the Elispot assay in healthy individuals and in patients with metastatic melanoma immunized with tyrosinase peptide. *Int J Cancer* 87:391-398.
- Lindemann M, Bohmer J, Zabel M, Grosse-Wilde H. 2003. ELISpot: a new tool for the detection of nickel sensitization. *Clin Exp Allergy* 33:992-998.
- Mantha S, Roizen M, Fleisher L, Thisted R, Foss J. 2000. Comparing methods of clinical measurement: reporting standards for bland and altman analysis. *Anesth Analg* 90:593-602.

- Mashishi T, Gray CM. 2002. The ELISPOT assay: an easily transferable method for measuring cellular responses and identifying T cell epitopes. *Clin Chem Lab Med* 40:903-910.
- Medhat A, Shehata M, Bucci K, Mohamed S, Dief AD, Badary S, Galal H, Nafeh M, King CL. 1998. Increased interleukin-4 and interleukin-5 production in response to *Schistosoma haematobium* adult worm antigens correlates with lack of reinfection after treatment. *J Infect Dis* 178:512-519.
- Merville P, Pouteil-Noble C, Wijdenes J, Potaux L, Touraine JL, Banchereau J. 1993. Detection of single cells secreting IFN-gamma, IL-6, and IL-10 in irreversibly rejected human kidney allografts, and their modulation by IL-2 and IL-4. *Transplantation* 55:639-646.
- Miyahira Y, Murata K, Rodriguez D, Rodriguez JR, Esteban M, Rodrigues MM, Zavala F. 1995. Quantification of antigen specific CD8+ T cells using an ELISPOT assay. *J Immunol Methods* 181:45-54.
- Mwau M, McMichael AJ, Hanke T. 2002. Design and validation of an enzyme-linked immunospot assay for use in clinical trials of candidate HIV vaccines. *SO - AIDS Res Hum Retroviruses* 2002 Jun 10;18(9):611-8. *AIDS Res Hum Retroviruses* 18:611-618.
- Nordstrom I, Ferrua B. 1992. Reverse ELISPOT assay for clonal analysis of cytokine production. II. Enumeration of interleukin-1-secreting cells by amplified (avidin-biotin anti-peroxidase) assay. *J Immunol Methods* 150:199-206.
- Otto K, Andersen MH, Eggert A, Keikavoussi P, Pedersen LO, Rath JC, Bock M, Brocker EB, Straten PT, Kampgen E, Becker JC. 2005. Lack of toxicity of therapy-induced T cell responses against the universal tumour antigen survivin. *Vaccine* 23(7):884-889.
- Pass HA, Schwarz SL, Wunderlich JR, Rosenberg SA. 1998. Immunization of patients with melanoma peptide vaccines: immunologic assessment using the ELISPOT assay. *Cancer J Sci Am* 4:316-323.
- Petersen P, Stockl D, Blaabjerg O, Pedersen B, Birkemose E, Thienpont L, Lassen J, Kjeldsen J. 1997. Graphical interpretation of analytical data from comparison of a field method with reference method by use of difference plots. *Clin Chem* 43:2039-2046.
- Pollock M, Jefferson S, Kane J, Lomax K, MacKinnon G, Winnard C. 1992. Method comparison--a different approach. *Ann Clin Biochem* 29:556-560.
- Romero P, Cerottini J, Waanders G. 1998. Novel methods to monitor antigen-specific cytotoxic T-cell responses in cancer immunotherapy. *Mol Med Today* 4:305-312.
- Sachs. 1978. *Angewandte Statistik. Statistische Methoden und ihre Anwendungen*. Springer Verlag, 1978.

- Schaed S, Klimek V, Panageas K, Musselli C, Butterworth L, Hwu W, Livingston P, Williams L, Lewis J, Houghton A, Chapman P. 2002. T-cell responses against tyrosinase 368-376(370D) peptide in HLA\*A0201+ melanoma patients: randomized trial comparing incomplete Freund's adjuvant, granulocyte macrophage colony-stimulating factor, and QS-21 as immunological adjuvants. *Clin Cancer Res* 8:967-972.
- Scheibenbogen C, Lee KH, Mayer S, Stevanovic S, Moebius U, Herr W, Rammensee HG, Keilholz U. 1997a. A sensitive ELISPOT assay for detection of CD8+ T lymphocytes specific for HLA class I-binding peptide epitopes derived from influenza proteins in the blood of healthy donors and melanoma patients. *Clin Cancer Res* 3:221-226.
- Scheibenbogen C, Lee KH, Stevanovic S, Witzens M, Willhauck M, Waldmann V, Naeher H, Rammensee HG, Keilholz U. 1997b. Analysis of the T cell response to tumor and viral peptide antigens by an IFNgamma-ELISPOT assay. *Int J Cancer* 71:932-936.
- Schmid D, Thieme M, Gary H, Reeves W. 1997a. Characterization of T cell responses to herpes simplex virus type 1 (HSV-1) and herpes simplex virus type 2 (HSV-2) using a TNF-beta ELISpot cytokine assay. *Arch Virol* 142:1659-1671.
- Schmid DS, Thieme ML, Gary HE, Jr., Reeves WC. 1997b. Characterization of T cell responses to herpes simplex virus type 1 (HSV-1) and herpes simplex virus type 2 (HSV-2) using a TNF-beta ELISpot cytokine assay. *Arch Virol* 142:1659-1671.
- Schmittel A, Keilholz U, Scheibenbogen C. 1997. Evaluation of the interferon-gamma ELISPOT-assay for quantification of peptide specific T lymphocytes from peripheral blood. *J Immunol Methods* 210:167-174.
- Schmittel A, Keilholz U, Thiel E, Scheibenbogen C. 2000. Quantification of tumor-specific T lymphocytes with the ELISPOT assay. *J Immunother* 23:289-295.
- Smith JG, Levin M, Vessey R, Chan IS, Hayward AR, Liu X, Kaufhold RM, Clair J, Chalikonda I, Chan C, Bernard M, Wang WW, Keller P, Caulfield MJ. 2003. Measurement of cell-mediated immunity with a Varicella-Zoster Virus-specific interferon-gamma ELISPOT assay: responses in an elderly population receiving a booster immunization. *J Med Virol* 70 Suppl 1:S38-41.
- Speiser D, Pittet M, Guillaume P, Lubenow N, Hoffman E, Cerottini J, Romero P. 2004. Ex vivo analysis of human antigen-specific CD8+ T-cell responses: quality assessment of fluorescent HLA-A2 multimer and interferon-gamma ELISPOT assays for patient immune monitoring. *J Immunother* 27:298-308.
- Tabidze IL, Lee FK, Tambe P, Rocha E, Larsen SA, Stoll BJ, St Louis ME, Nahmias AJ. 1999. Enzyme-linked immunospot assay for the diagnosis of active *Treponema pallidum* infection during the various stages of syphilis. *Sex Transm Dis* 26:426-430.

- Tambur A, Ben-Ari Z, Herrera N, Klein T, Michowiz R, Mor E. 2005. Donor-specific hyporesponsiveness in ELISPOT assay is associated with early recurrence of hepatitis C in liver transplant recipients [In Process Citation]. *Hum Immunol* 66:21-27
- Thumann P, Moc I, Humrich J, Berger TG, Schultz ES, Schuler G, Jenne L. 2003. Antigen loading of dendritic cells with whole tumor cell preparations. *J Immunol Methods* 277:1-16.
- van der Burg SH, Rensing ME, Kwappenberg KM, de Jong A, Straathof K, de Jong J, Geluk A, van Meijgaarden KE, Franken KL, Ottenhoff TH, Fleuren GJ, Kenter G, Melief CJ, Offringa R. 2001. Natural T-helper immunity against human papillomavirus type 16 (HPV16) E7-derived peptide epitopes in patients with HPV16-positive cervical lesions: identification of 3 human leukocyte antigen class II-restricted epitopes. *Int J Cancer* 91:612-618.
- Wang F, Bade E, Kuniyoshi C, Spears L, Jeffery G, Marty V, Groshen S, Weber J. 1999. Phase I trial of a MART-1 peptide vaccine with incomplete Freund's adjuvant for resected high-risk melanoma. *Clin Cancer Res* 5:2756-2765.
- Yang J, Lemas VM, Flinn IW, Krone C, Ambinder RF. 2000. Application of the ELISPOT assay to the characterization of CD8(+) responses to Epstein-Barr virus antigens. *Blood* 95:241-248.

## **Danksagung**

Das Thema der Arbeit geht auf Frau Prof. Dr. med. E.-B. Bröcker und Frau Dr. med. S. Schäd zurück. Ihnen möchte ich für die großzügige Unterstützung, die vielen Anregungen und die gute Zusammenarbeit danken.

Frau Dr. Schäd danke ich insbesondere für die Einführung in die Thematik und das Erlernen der Elispot-Methode, für den regen Gedankenaustausch, sowie für das intensive Korrekturlesen der Arbeit.

Mein Dank gilt weiterhin Herrn Dr. W. Malkusch für die intensive Unterstützung bei Fragestellungen zur Technik und Bildanalyse.

Danken möchte ich auch besonders Herrn Dr. U. Mäder, der mir mit seinem mathematisch-statistischen Rat immer zur Seite stand.

Herrn Prof. Dr. med. H. J. Girschick danke ich für die Unterstützung der Arbeit in seinem Labor.

Ganz besonders danke ich auch Herrn Dipl.-Inf. Thomas Schaal für seine Unterstützung bei Fragen der Bildbearbeitung und Formatierung dieser Arbeit.

## Curriculum vitae

**Name:** Alexandra Beata Vogtmann  
**Geburtsdatum:** 14.07.1976  
**Geburtsort:** Laurahütte (Siemianowice Slaskie, PL)  
**Staatsangehörigkeit:** deutsch

---

**Eltern:** Marian Vogtmann, geb. 22.05.1947  
Irene Babis-Vogtmann, geb. 21.04.1947  
**Familienstand:** ledig  
**Konfession:** römisch-katholisch

---

### Schulausbildung:

1983 - 1987 Grundschule in Laurahütte  
1987 - 1988 Maria-Theresia-Volksschule in Günzburg  
1988 - 1997 Humanistisches und neusprachliches  
Kronberg Gymnasium Aschaffenburg

### Schulabschluss:

1997 Allgemeine Hochschulreife, Note 1,7

---

### Studium:

10/1997 - 02/1998 Technische Universität Darmstadt  
*Studiengang: Psychologie*  
05/1998 - 11/2004 Julius-Maximilians-Universität Würzburg  
*Studiengang: Humanmedizin*

Examina Physikum April 2000, Note 2,7  
1. Staatsexamen März 2001, Note 2,0  
2. Staatsexamen September 2003, Note 2,0  
3. Staatsexamen November 2004, Note 1,0

Praktische Ausbildung Juliusspital Würzburg, Chirurgie  
Klinikum Aschaffenburg, Innere Medizin  
Universitätsklinikum Würzburg, Dermatologie

Approbation November 2004  
Dissertation

Beginn der vorliegenden Arbeit Oktober 2001

---

**berufliche Tätigkeit:**

06 / 2005 Beginn der Tätigkeit als Assistenzärztin in der  
Medizinischen Klinik I des Klinikums Aschaffenburg,  
akademisches Lehrkrankenhaus der Universität Würzburg

A. Ob | —