

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Dutruef, C; Thole, J; Geels, M; Mollenkopf, HJ; Ottenhoff, T; Guzman, CA; Fletcher, HA; Leroy, O; Kaufmann, SH (2014) TRANSVAC workshop on standardisation and harmonisation of analytical platforms for HIV, TB and malaria vaccines: 'How can big data help?'. *Vaccine*. ISSN 0264-410X DOI: <https://doi.org/10.1016/j.vaccine.2014.06.014>

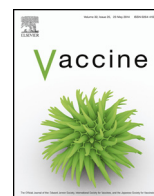
Downloaded from: <http://researchonline.lshtm.ac.uk/1805440/>

DOI: [10.1016/j.vaccine.2014.06.014](https://doi.org/10.1016/j.vaccine.2014.06.014)

Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: <http://creativecommons.org/licenses/by-nc-nd/2.5/>



Brief report

TRANSVAC workshop on standardisation and harmonisation of analytical platforms for HIV, TB and malaria vaccines: 'How can big data help?'



Céline Dutruel^a, Jelle Thole^b, Mark Geels^a, Hans-Joachim Mollenkopf^c, Tom Ottenhoff^d, Carlos A. Guzman^e, Helen A. Fletcher^f, Odile Leroy^a, Stefan H.E. Kaufmann^{g,*}

^a European Vaccine Initiative, Heidelberg University Hospital, Im Neuenheimer Feld 326, 69120 Heidelberg, Germany

^b Tuberculosis Vaccine Initiative, Runderweg 6, 8219 PK Lelystad, The Netherlands

^c Core Facility Microarray, Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany

^d Department of Infectious Diseases, Leiden University Medical Center, Bldg. 1, Rm # C-05-44, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

^e Department of Vaccinology and Applied Microbiology, Helmholtz Centre for Infection Research, Inhoffenstraße 7, 38124 Braunschweig, Germany

^f Immunology and Infection, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom

^g Department of Immunology, Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany

ARTICLE INFO

Article history:

Received 15 January 2014

Received in revised form 26 May 2014

Accepted 6 June 2014

Available online 17 June 2014

Keywords:

Vaccination

Standardisation

Harmonisation

Molecular analyses

Biomarkers

Malaria

Tuberculosis

HIV

TRANSVAC

ABSTRACT

High-throughput analyses of RNA and protein expression are increasingly used for better understanding of vaccine-induced immunity and protection against infectious disease. With an increasing number of vaccine candidates in clinical development, it is timely to consider standardisation and harmonisation of sample collection, storage and analysis to ensure results of highest quality from these precious samples. These challenges were discussed by a group of international experts during a workshop organised by TRANSVAC, a European Commission-funded Research Infrastructure project. The main conclusions were: Platforms are rarely standardised for use in preclinical and clinical studies. Coordinated efforts should continue to harmonise the experimental set up of these studies, as well as the establishment of internal standards and controls. This will ensure comparability, efficiency and feasibility of the global analyses performed on preclinical and clinical data sets.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Abbreviations: dcRT-MLPA, dual colour reverse transcription multiplex ligation-dependent probe amplification; EC, European Commission; ELISA, enzyme linked immunosorbent assay; FP, Framework Programme; HIV, human immunodeficiency virus; HVTN, HIV Vaccine Trials Network; IMI, innovative medicines initiative; PBMC, peripheral blood mononuclear cell; QA, quality assurance; QC, quality control; RNA, ribonucleic acid; RT-PCR, real-time polymerase chain reaction; SOP, standard operating procedures; TB, tuberculosis.

* Corresponding author. Tel.: +49 3028 460 500.

E-mail addresses: celine.dutruel@euvaccine.eu (C. Dutruel), jelle.thole@intravacc.nl (J. Thole), mjgeels@gmail.com (M. Geels), mollenkopf@mpiib-berlin.mpg.de (H.-J. Mollenkopf), t.h.m.ottenhoff@lumc.nl (T. Ottenhoff), carlos.guzman@helmholtz-hzi.de (C.A. Guzman), helen.fletcher@shtm.ac.uk (H.A. Fletcher), odile.leroy@euvaccine.eu (O. Leroy), kaufmann@mpiib-berlin.mpg.de (S.H.E. Kaufmann).

<http://dx.doi.org/10.1016/j.vaccine.2014.06.014>

0264-410X/© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

1. Introduction

Global molecular analyses are exploited to enhance our understanding of novel vaccination strategies. High-throughput technologies, including microarray analyses and RNA deep sequencing, allow genome-wide profiling of gene expression within different study groups. Similarly, targeted assays enable study of the expression of a dedicated number of genes [e.g. dual colour reverse transcription multiplex ligation-dependent probe amplification (dcRT-MLPA) assay], cell-expressed molecules (e.g. flow cytometry) or secreted molecules (multiplex assays).

Expectations of data output from these analyses in vaccine trials are high, and it is hoped that through the systematic analysis of biomarkers using modern bioassays, predictive biomarkers, which can be used as (surrogate) markers of clinical endpoints or of adverse events, can be identified. In tuberculosis (TB), human immunodeficiency virus (HIV), and malaria, few approaches are on the pathway for use with clinical trial samples for prediction of both vaccine efficacy and vaccine safety. However, molecular analytical tools are providing first hints regarding mechanisms underlying protection against, or susceptibility to, developing clinical disease [1–3]. Since there are now a number of vaccine candidates in phase II/III clinical trials in the TB, HIV, and malaria arenas, it is timely to consider standardisation and harmonisation of sample collection, storage and molecular analysis to ensure highest quality data from these precious samples.

In order to discuss these challenges a workshop was organised by TRANSVAC, a European Commission (EC)-funded project coordinated by the European Vaccine Initiative. The aim of the workshop was to define and implement a process supporting the harmonisation of operational procedures for the profiling and the assessment of novel vaccine candidates, novel vaccine formulations, and/or novel routes of administration. Through internal research activities in the field of HIV, TB, and malaria, and through the supply of services to 24 projects, including free access to adjuvants, animal models, microarray analysis, and assays/standards, the TRANSVAC partners have contributed to harmonisation of protocols. These efforts, which took place between 2009 and 2013, were discussed at the TRANSVAC workshop.

2. Overview from data sets obtained with different technologies, diseases and species, and in different clinical settings

2.1. Genome-wide high-throughput technologies

To obtain meaningful data sets from preclinical studies and clinical trials, standardisation and harmonisation of sample collection, storage and analysis are crucial. Results performed with three genome-wide high-throughput technologies (Agilent Technologies and Affymetrix transcriptome platforms, as well as Illumina sequencing platform) were presented [4,5]. While sample collection and pre-processing of the samples (e.g. RNA isolation, labelling for microarray analysis and library generation for sequencing) are well standardised, analysis was confounded by different influences, including the nonhuman primate sub-species analysed, the health history of study participants, and by differences in the sources of RNA (e.g. cell-free nucleic acids and platelet RNA, both derived from different types of blood cells).

It was concluded that essential factors for studies involving microarrays are (i) group sizes, (ii) timepoints of measurement (including multiple pre-vaccination time points to account for inter-individual variation), (iii) strength of vaccine-induced responses, (iv) nature of test samples, and (v) quality of test samples.

Previous studies have found that, depending on sequencing depth, next-generation sequencing platforms can be more comprehensive than microarrays in detecting expression differences and have no hybridisation bias [6,7], but are computationally more complex and time consuming. Nevertheless, computational bioinformatics' analyses are essential for both techniques to obtain meaningful data and to compare data sets, and can best be embedded at the research group level [8,9].

2.2. Other technologies

Global analysis of metabolites (metabolomics) in serum has proven its value for biomarker design in infectious diseases including monitoring of vaccine trials [7].

Flow cytometric analysis and/or mass cytometric analysis of cells or cell-bound proteins can be used as predictive biomarkers for disease outcome and response to immune interventions [10]. These approaches seem to be more powerful than conventional methods, such as ELISA and Luminex, with key features like a short sample processing time, low blood amounts required per condition to be tested, the possibility to process both stimulated or non-stimulated samples, and the use of fresh samples which reduces the artefacts and loss of sensitivity due to cryopreservation. Important issues to guarantee reliability of the obtained data are standardisation of sample preparation, transport and storage, inter-test variation (occurring when large numbers of samples are processed by a single operator on a single day), data acquisition, and appropriate quality controls (QCs) (e.g. acceptable percentage of dead cells, minimum number of analysed events, reference controls). In the field of cancer immunotherapy, harmonisation and standardisation of T-cell immunoassays (e.g. ELISpot and intracellular cytokine staining) has proven to be feasible on an international scale with great success [11].

Growth inhibition assays are increasingly used in TB and malaria. For TB, whole blood or PBMC-based tests utilising a liquid culture system for detection of mycobacterial growth have shown promise and are currently being assessed for use in early phase vaccine clinical trials [12,13].

As an alternative to array-based platforms, assays have been designed that offer specific, robust, affordable and practical bio-profiling platforms. The dcRT-MLPA assay is a RT-PCR-based gene expression profiling method, which represents a valid alternative to perform intermediate sized multiplex screens [1,3] once a tailored signature has been composed, e.g., based on information from unbiased genome-wide expression analysis. The assay setup ensures high assay sensitivity and avoids the limitations of multiplex PCR and the costly aspects of genome-wide platforms such as microarrays and RNA sequencing.

2.3. Integration of epidemiological data

It is becoming increasingly obvious that type of samples used (e.g. whole blood, PBMC, serum, plasma and urine), age of the individuals, or environmental factors (e.g. the circadian rhythm of the subjects including the number of sleep hours) can have a great impact on host responses [14]. It is thus important to carefully monitor epidemiological data from clinical trial study participants to draw adequate conclusions, when analysing the data.

In the context of clinical trials, systems biology combines clinical and epidemiological data with all transcriptional, proteomic, metabolomic and immunological data gathered [8,9,15–19]. There are three interlinking components of systems biology: experimental design, biological questions and data analysis. This results in a data set that ultimately needs to be validated before it can be usefully applied. Tools are available that can greatly reduce data

complexity and help in the identification of biomarkers, but oversimplification may lead to loss of insight into pathomechanisms.

A major bottleneck remains the difficulty to sustain a highly controlled environment in phase I clinical trials, during the time period between vaccination and the expected “operation” time of the vaccine. Moreover, to fully correct for all the parameters influencing the data, sampling schedules including a high number of critically chosen samples and time points are needed, but are frequently ignored due to time and cost restrictions. A trade-off thus has to be found between the amount of data that can be obtained and the means and know-how available to analyse the collected data.

2.4. Comparison of data from different preclinical/clinical trials

A number of EC Framework Programme (FP) 6 and FP7 projects (i.e. TBVAC/NEWTBVAC, ADITEC, Euronut41, OPTIMALVAC and EMVDA), and the IMI project BioVacSafe have contributed to standardisation of different protocols and SOPs, in order to allow comparison of readouts between different clinical trial sites. While strict reporting forms are well advanced [20–22], bottlenecks are time frame differences and investigator-specific protocols.

A different approach is to centralise all immunological readouts. The HIV Vaccine Trials Network (HVTN, Dr. Julie McElrath) is the quintessential example of a centralised infrastructure driving and executing the analysis of vaccine-induced immune responses in large clinical trials. HVTN has centralised use of qualified and validated immune assays, of common reagents, and of archived specimens, as well as collaborations and infrastructures including advanced planning. A centralised lead laboratory is responsible for quality assurance (QA)/QC and the repository of samples, while specialised working groups take care of protocols, support and QC of specimen [22]. Notable trials that were evaluated by HVTN were the HIV-1 STEP and RV144 trials [23,24].

3. Recommendations and conclusions

Only few global analysis platforms are fully standardised to inform and allow informative use in preclinical studies and clinical trials through which licensure could be obtained. Coordinated efforts between different disease networks should continue to achieve standardisation of immunological and global platforms that will allow their effective use in a clinical setting, their use for biomarker discovery and validation, and their use in generating data sets that can be compared between different platforms and across different preclinical settings and/or different clinical trials.

The main challenges to be overcome when performing global analyses can be grouped into the following:

- I. Definition of study group sizes and numbers in order to compare studies.
- II. Recommendations for sample collection and storage for future analysis wherever possible according to generally agreed-upon recommendations.
- III. Understanding of the influence of the environment on host responses. Not all variables can be taken into account when performing clinical studies and it is important to identify the parameters of greatest influence and to harmonise their monitoring.
- IV. Establishment of internal standards and controls to ensure comparability, efficiency and feasibility of molecular and immunological profiling.

These challenges require international collaboration, which could be led by the Systems Biology Platform of the European

Vaccine Research and Development Infrastructure, in association with regulatory agencies and representatives of the World Health Organisation. The process of harmonising methodology, sample and data collection, and the analysis of data will benefit from previous experiences in ADITEC and BIOVACSAFE European projects, together with the NIAID-sponsored Systems Biology for Infectious Diseases Research Program. The working parties should agree on core recommendations and a strategic action plan to address these priorities. If funding for European Vaccine Research and Development Infrastructure materialises in 2015, a pilot phase will be launched for structuring global analyses of infectious diseases with high public health importance, such as AIDS, tuberculosis, malaria, and influenza.

Acknowledgements

We would like to thank speakers at the Global Analyses Platforms Workshop: Sabin Bhujy, Carlos A. Guzman, Stefan H.E. Kaufmann, Helen Fletcher, David Lewis, Julie McElrath, Hans-Joachim Mollenkopf, Tom Ottenhoff, Steven Smith, Thomas Stempf, Robert van den Berg, and Frank Verreck, without whom this manuscript could not have been written. We also thank the TRANSVAC consortium (www.transvac.org) as well as Regitze Thøgersen, TRANSVAC project leader, for her help in the organisation of the workshop and management of the project. The workshop was supported by the EU FP7 project TRANSVAC (FP7-INFRASTRUCTURES-2008-228403). This work has received support from the EU/EFPIA Innovative Medicines Initiative Joint Undertaking ([BioVacSafe] grant n° [115308], the EU FP6 funded project EMVDA (LSHP-CT-2006-037506), and the EU FP7 funded projects NEWTBVAC (FP7-HEALTH-2009-241745), ADITEC (FP7-HEALTH-2011-280873) and EeURONEUT-41 (FP7-HEALTH-2007-201038). This publication reflects only the authors' views and the European Union is not liable for any use that may be made of the information contained herein.

Conflict of interest: None declared.

References

- [1] Joosten SA, Goeman JJ, Sutherland JS, Opmeer L, de Boer KG, Jacobsen M, et al. Identification of biomarkers for tuberculosis disease using a novel dual-color RT-MLPA assay. *Genes Immun* 2012;13:71–82.
- [2] Thakur A, Pedersen LE, Jungersen G. Immune markers and correlates of protection for vaccine induced immune responses. *Vaccine* 2012;30:4907–20.
- [3] Sutherland JS, Hill PC, Adetifa IM, de Jong BC, Donkor S, Joosten SA, et al. Identification of probable early-onset biomarkers for tuberculosis disease progression. *PLoS One* 2011;6:e25230.
- [4] Maertzdorf J, Weiner J, Mollenkopf H-J, Bauer T, Prasse A, Müller-Quernheim J, et al. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. *Proc Natl Acad Sci U S A* 2012;109:7853–8.
- [5] Maertzdorf J, Reipsilber D, Parida SK, Stanley K, Roberts T, Black G, et al. Human gene expression profiles of susceptibility and resistance in tuberculosis. *Genes Immun* 2011;12:15–22.
- [6] Malone JH, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* 2011;9:34.
- [7] Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 2008;18:1509–17.
- [8] Maertzdorf J, Weiner J, Kaufmann SHE. Enabling biomarkers for tuberculosis control. *Int J Tuberc Lung Dis* 2012;16:1140–8.
- [9] Weiner J, Maertzdorf J, Kaufmann SHE. The dual role of biomarkers for understanding basic principles and devising novel intervention strategies in tuberculosis. *Ann N Y Acad Sci* 2013;1283:22–9.
- [10] Zaritskaya L, Shurin MR, Sayers TJ, Malyguine AM. New flow cytometric assays for monitoring cell-mediated cytotoxicity. *Expert Rev Vaccines* 2010;9:601–16.
- [11] van der Burg SH, Kalos M, Gouttefangeas C, Janetzki S, Ottensmeier C, Welters MJ, et al. Harmonization of immune biomarker assays for clinical studies. *Sci Transl Med* 2011;3, 108ps44.
- [12] Goodman AL, Draper SJ. Blood-stage malaria vaccines – recent progress and future challenges. *Ann Trop Med Parasitol* 2010;104:189–211.
- [13] Fletcher HA, Tanner R, Wallis RS, Meyer J, Manjaly Z-R, Harris S, et al. Inhibition of mycobacterial growth in vitro following primary but not secondary vaccination with *Mycobacterium bovis* BCG. *Clin Vaccine Immunol* 2013;20(11):1683–9.

- [14] Ackermann K, Revell VL, Lao O, Rombouts EJ, Skene DJ, Kayser M. Diurnal rhythms in blood cell populations and the effect of acute sleep deprivation in healthy young men. *Sleep* 2012;35:933–40.
- [15] van Mechelen I, Smilde A. A generic linked-mode decomposition model for data fusion. *Chemom Intell Lab* 2010;104(1):83–94.
- [16] van Mechelen I, Smilde A. Comparability problems in the analysis of multiway data. *Chemom Intell Lab* 2011;106:2–11.
- [17] van Deun K, Smilde AK, van der Werf MJ, Kiers HA, van Mechelen I. A structured overview of simultaneous component based data integration. *BMC Bioinform* 2009;10:246.
- [18] Bylesjö M, Eriksson D, Kusano M, Moritz T, Trygg J. Data integration in plant biology: the O2PLS method for combined modeling of transcript and metabolite data. *Plant J* 2007;52:1181–91.
- [19] Hanafi M, Kiers H. Analysis of K sets of data, with differential emphasis on agreement between and within sets. *Comput Stat Data Anal* 2006;51: 1491–508.
- [20] Hanekom WA, Dockrell HM, Ottenhoff THM, Doherty TM, Fletcher H, McShane H, et al. Immunological outcomes of new tuberculosis vaccine trials: WHO panel recommendations. *PLoS Med* 2008;5:e145.
- [21] Smith SG, Joosten SA, Verscheure V, Pathan AA, McShane H, Ottenhoff THM, et al. Identification of major factors influencing ELISpot-based monitoring of cellular responses to antigens from *Mycobacterium tuberculosis*. *PLoS One* 2009;4:e7972.
- [22] Britten CM, Janetzki S, Butterfield LH, Ferrari G, Gouttefangeas C, Huber C, et al. T cell assays and MIATA: the essential minimum for maximum impact. *Immunity* 2012;37:1–2.
- [23] Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N Engl J Med* 2012;366:1275–86.
- [24] McElrath MJ, de Rosa SC, Moodie Z, Dubey S, Kierstead L, Janes H, et al. HIV-1 vaccine-induced immunity in the test-of-concept step study: a case-cohort analysis. *Lancet* 2008;372:1894–905.