

Quality Control for the Molecular Diagnosis of Toxoplasmosis

Emmanuelle Varlet-Marie, Yvon Sterkers and Patrick Bastien
Centre Hospitalier Universitaire (C.H.U.) of Montpellier / University of Montpellier 1,
Laboratoire de Parasitologie-Mycologie, Laboratory Associated (Pôle "Biologie
Moléculaire") to the National Reference Centre for Toxoplasmosis,
Montpellier
France

1. Introduction

Toxoplasmosis is an endemic parasitic disease due to the protozoon *Toxoplasma gondii*. The definitive host is the cat, in which the parasite develops in the intestinal epithelium, before being eliminated as oocysts in the faeces. When oocysts, for example in contaminated soil, are ingested by humans, other mammals or birds, they pass through the stomach and excyst in the intestine; the released sporozoites invade the intestinal wall and give rise to generalized infection (Remington & Desmonts, 1995; Wong & Remington, 1993). Two tissue stages are important in pathogenesis: tachyzoites and bradyzoites. Initially, the infection consists mainly of rapidly dividing tachyzoites. Then, with developing immunity, persistent tissue cysts containing hundreds of slowly multiplying bradyzoites develop, especially in the muscle and brain. This ability to live inside cells without destroying them allows evasion of host immunity and the development of a chronic stage of infection, lasting for years. Intermediate hosts become infected either by ingesting soil, water or foodstuff contaminated with cat faeces or by ingesting raw or undercooked meat containing tissue cysts.

T. gondii is responsible for generally benign infections except when the disease occurs in pregnant women (congenital toxoplasmosis) or in immunocompromised individuals, such as human immunodeficiency virus-positive or grafted patients, in which cases the vital prognosis may be involved. In certain countries of Europe, including France, toxoplasmosis is regarded as a serious health problem. In France, the prevalence of acquired toxoplasmosis in adults is 44%, and the estimated yearly incidence of contamination in women during pregnancy and of congenital toxoplasmosis are high, respectively, 6-7/1000 and 0.1% of births (Berger et al., 2008, King et al., 2008). Prevention of congenital toxoplasmosis (CT), including prenatal diagnosis (PND), has become a national policy in France since 1978 (Thulliez, 1992). Serological screening and follow-up is established for non-immunized pregnant women, and associated with monthly ultrasound examinations. PND is proposed in case of seroconversion between 6 to 38 weeks of amenorrhea. For this, PCR-based molecular diagnostic tests using amniotic fluid, introduced in the early 90's, have rapidly eliminated the need for cordocentesis, and have competed with more classical biological methods. Wherever it has been implemented, PND of CT has clearly improved the prognosis and outcome of infected children (reviewed in (Bastien, 2002)).

Because the host immune status plays an important role in the clinically patent expression of toxoplasmosis (Luft & Remington, 1992), particularly when CD4-positive cells reach counts $<200/\mu\text{l}$ (Nissapatorn & Abdullah, 2004), toxoplasmosis is also one of the most common complications in immunocompromised hosts (HIV-infected and transplanted patients). Screening for toxoplasmosis in hematology patients has therefore become the most important request for *Toxoplasma*-PCR in Parasitology departments today.

The French national prevention policy was recently reinforced by the creation of a National Reference Centre for Toxoplasmosis (CNRT) (<http://www.chu-reims.fr/professionnels/cnr-toxoplasme-1/>), which includes a working group for molecular biology, of which the objectives are, among others, the improvement and standardization of the molecular diagnosis of toxoplasmosis at the national level, and of which one of us (PB) is the coordinator. In France, the PND of CT is realized essentially in University hospitals, as well as two large private biological diagnosis centers. Not all University hospital centers perform such testing, as centers and practitioners involved in this diagnosis need an official authorization for this from the national health authorities, which is granted for five years.

In total, like for many infectious diseases, PCR-based diagnostic tests have become essential in the diagnosis of toxoplasmosis. However, also like in many other fields of clinical biology, these assays suffer from a great lack of standardization, as the vast majority of centers performing this diagnosis use 'in house' or laboratory-developed PCR assays; i.e. that have been set up independently in each laboratory using different DNA targets, and customized primers and reaction conditions.. This leads to a great heterogeneity between laboratories. On top of this, 'in house' methods can largely differ at any step during the diagnostic process, such as the extraction method, the number of PCR tubes used for diagnosis, the inclusion of an internal control for detecting inhibitions of the reaction, etc.(Sterkers et al., 2009). All this may be the source of considerable interlaboratory variations in the performances of the assays, influencing the quality of the diagnosis and hampering valuable comparisons of data among different centers. Previous studies have highlighted the lack of homogeneity and performance in European countries and underlined the need for guidelines (Pelloux et al., 1998, Kaiser et al., 2007, Bastien et al., 2007, Sterkers et al., 2009, Sterkers et al., 2010). In view of this heterogeneity, standardization of PCR methods and practices has become a strong wish of both the health authorities and the community of clinical microbiologists. Such a standardization should in turn lead to improvement of the diagnosis of CT at a more global level.

However, in view of the high diversity of methods and equipments available (see below), a true standardization of methods appears impossible, rendering external quality assessments (EQAs) the only way to homogenize the level of performances of this diagnosis among laboratories. The necessity of accreditation of the clinical laboratories according to a Quality management system added weight to these requirements. Indeed, this should include "comprehensive validation of each assay using the appropriate specimen matrix, daily quality control measures, and the enrolment in an appropriate quality assurance program for each assay to ensure continued performance, the whole being associated to the recording of results to document and address any inadequacies in test performance long-term" (Beld et al., 2007). EQAs for the molecular diagnosis of CT have been the subject of a number of small-scale studies in Europe (Guy et al., 1996, Pelloux et al., 1998) before being implemented at a national (Bastien et al., 2007) or European (Kaiser et al., 2007) scale, from

2002 and 2004, respectively, onwards. Toxoplasmosis is, to our knowledge, the only parasitic infection for which EQAs for molecular diagnosis exist.

In this chapter, we shall describe the experience of eight years of implementation of the French national EQA; we shall also discuss the interest, specificities, protocols, qualitative and quantitative results, as well as the perspectives, of EQAs for the molecular diagnosis of toxoplasmosis.

2. Eight years of national EQAs

In order to implement the harmonization of prenatal diagnosis of CT in France, an early initiative for quality assurance in the molecular prenatal diagnosis of toxoplasmosis was launched by the French association of hospital practitioners and teachers in Parasitology-Mycology (ANOFEL) in 2002. Briefly, on a yearly basis, a panel of positive and negative amniotic fluid samples prepared in Montpellier was sent blinded to participating centers for PCR testing, allowing each center to assess and follow its own performances in the molecular detection of CT (Bastien et al., 2007). This EQA was then continued by the working group for molecular diagnosis of the CNRT.

An annual national survey was associated to the EQA in order to assess the diversity and evolution of methods and practices in this molecular diagnosis in France. The questionnaires included 11 queries, concerning what was considered as the most critical points of the PCR process (Sterkers et al., 2009). All answers to queries had to be given considering the routine practice of prenatal diagnosis of CT, and not what could have been realized for the EQA only.

On top of these national 'core' EQAs, more limited and specific ('scientific') EQAs aiming at refining comparative studies between PCR methods were implemented at a smaller scale. A multicentric prospective study, involving eight laboratories proficient in the molecular prenatal diagnosis of toxoplasmosis, was a first step towards the harmonization of this diagnosis among university hospitals in France (Sterkers et al., 2010). It aimed at comparing the analytical performances of different PCR protocols used for *Toxoplasma* detection, and was reproduced over two consecutive years. Each centre extracted the same concentrated *T. gondii* suspension and tested serial dilutions of the extracted DNA using its own assays. Since all reactions are not positive when the concentration of pathogens gets to the sensitivity limit of the method (Chabbert et al., 2004; Sterkers et al., 2010), original 'performance scores' were defined, taking into account the proportion of positive reactions over the number of reactions performed, which appeared as a good indicator of sensitivity. Differences in analytical sensitivities were observed between assays, particularly at low parasite concentrations (≤ 2 *T. gondii* genomes per reaction tube, i.e. ≤ 10 parasites/mL depending on the different pre-analytical protocols used in the participating centre), with 'performance scores' varying by a 20-fold factor among laboratories. Our data stressed the fact that differences do exist in the performances of molecular assays in spite of expertise in the matter. Finally, these studies allowed proposing a PCR sensitivity threshold to diagnostic laboratories as a minimal objective to reach for an optimal molecular diagnosis of toxoplasmosis.

Another multicentric prospective study aimed at comparing the analytical performances of different PCR protocols used for *Toxoplasma* detection, this time at very low concentrations (2.5 and 1 T/mL). Each of eight centers received 3 lyophilized samples of a *T. gondii* suspension at both concentrations, and extracted and tested each sample. This (i) confirmed

the differences in analytical sensitivities observed between assays, and (ii) allowed to experimentally assess the minimum number of samples to be tested at these very low concentrations, in relation to the heterogeneity of distribution linked to the Poisson law of large numbers (Varlet et al. article in preparation).

3. Specificities of EQAs for *Toxoplasma* molecular detection

The specificities of this EQA are related to (i) the fragility of *T. gondii* tachyzoites, and (ii) the low parasitic loads frequently encountered in protozoal infections.

We have observed that tachyzoites are extremely and rapidly subject to degradation at ambient temperature, even in nutritive media such as in vitro cultivation media or biological fluids (unpublished). This may occur during transport to participating centers or even during manipulations on the bench. First, this implies that all manipulations be carried out on ice and any temporary storage of the samples made at +4°C. Second, this gave a strong impulse towards freeze drying as the most adapted preparation method for EQAs (see below).

The demonstration of low parasite loads in a large proportion of infected AF samples (<10 tachyzoites per mL) (Costa et al., 2001; Romand et al., 2004) makes it necessary to test sensitivity with low *Toxoplasma* concentrations. Moreover, this choice allows a finer and more stringent assessment of the performances of the methods used by each laboratory year by year, as diagnostic methods for pathogens are particularly fallible with low concentrations of pathogens in the biological sample (Chernesky et al., 2002; Kaiser et al., 2007; Lachaud et al., 2001; Lachaud et al., 2002; Pelloux et al., 1998). It is noteworthy that these considerations probably apply to most other parasitic diseases.

4. Different protocols for EQA of *Toxoplasma*-PCR

A large number of variables have to be taken in consideration when preparing such an EQA. All of them must be subject to a choice made on both scientific, technical and practical grounds.

The matrix has a profound impact upon the sensitivity of the PCR assay. Indeed, the presence of inhibitors such as hemoglobin, blood proteins, anticoagulants etc. (Al-Soud & Radstrom, 2001) is obviously more a problem when analyzing blood-related products than more 'simple' biological fluids such as amniotic fluid, cerebrospinal fluid or aqueous humor. Also of note, the total amount of DNA in the template preparation, which includes DNA from the pathogen and its host, is also much higher in the former than in the latter. This, in turn, directly influences the amount of free Mg²⁺ ions in the reaction mix, and hence the enzymatic activity of the DNA polymerase (Bastien et al., 2008). Thus, the observed analytical sensitivity of the assay is somehow proportional to the volume of template DNA input in the PCR reaction (Burggraf et al., 2005) and high concentrations of human DNA do decrease the sensitivity of molecular assays designed to detect infectious pathogens. Another impact of the sample matrix is its effect on extraction efficiency. Indeed, it is well-known that the paucity of cells in a biological fluid leads to inefficient DNA recovery, due to the absence of "carrier-DNA" effect during extraction. Therefore, it is not recommended to use a non-cellular fluid such as saline (NaCl 0.9%) or PBS or culture medium as a matrix. Actually, the working group for molecular diagnosis of the CNRT did use saline but with high concentrations of the parasite (> 10⁵/mL), in order to precisely avoid this problem.

QCMD, a 'not for profit' European organization dedicated to EQAs in molecular diagnosis in microbiology, is also using plasma as a matrix, seemingly successfully. Yet, we believe further testing is required to assess the extraction efficiency of low concentrations of *T. gondii* in plasma.

All EQAs for *Toxoplasma*-PCR published to date used panels consisting of samples of human amniotic fluid (Bastien et al., 2007, Kaiser et al., 2007). Although this avoids tackling with the problem of PCR inhibition, it allows good extraction efficiency and perfect mimicry of a diagnostic sample. In spite of the fact that amniotic fluid drawn from hydramnios is often naturally diluted, hence not ideal because pauci-cellular, it constitutes the most common source of matrix, and is collected after informed consent has been obtained. There is no consensus as to whether the patient should be seronegative for toxoplasmosis or seropositive with proven sound immunity due to past toxoplasmosis.

Samples from several patients are pooled in order to constitute a homogeneous matrix. The presence of blood should be avoided except if intentionally added to induce inhibition. Of note is the fact that we also tested bovine amniotic fluid, which proved rather satisfactory, but was abandoned because (a) it often presented bacterial contaminations and (b) it impeded the detection of human beta-globin by participants (as an inhibition control).

The parasitic material to be tested may be previously extracted DNA or whole *T. gondii* tachyzoites. DNA can be used as such in its own extraction buffer. However, our experience demonstrated that the DNA extraction process and the PCR assay both contribute to the performance of the whole diagnostic method (Sterkers et al., 2010); thus, this form of EQA does not allow testing the DNA extraction method, and this can even lead to wrong assessments of the method used in the participating laboratories (Molecular biology working group of the CNRT, unpublished data).

Therefore, the use of whole tachyzoites is preferred. Yet, these can originate from different sources and be sent under different forms.

Between 2002 and 2004, our test material was based upon naturally infected samples kept at -20°C after primary diagnosis. Known PCR-positive and PCR-negative amniotic fluids were pooled separately, and the pools were tested for the presence of bacterial contamination and by *Toxoplasma*-PCR. The reference PCR assay used to control negative samples allowed the detection of 0.5-1 parasite/mL (Chabbert et al., 2004). The negative pool was used for making negative control samples, and also for diluting the positive pool in order to produce samples with definite *Toxoplasma* concentrations. The method was simple and had the advantages of (a) testing naturally infecting strains (b) sending non-infectious samples (thawed parasites being dead). Drawbacks included an increased risk of bacterial contamination (leading to DNA degradation during transport), a reduced accuracy in the quantification of parasites (parasite concentrations in the positive pool were estimated by quantitative PCR in two or three different laboratories before dilution and distribution), as well as, later, ethical problems in France with respect to this type of samples (even though signed consent was obtained from all women from whom the samples were drawn).

After that, the test material was based upon the classical method of seeding negative amniotic fluids with live *T. gondii* tachyzoites. Between 2005 and 2007, parasites were collected from ascitic fluid containing tachyzoites drawn from mouse infected with the RH strain of *T. gondii*. These were added at different concentrations into thawed amniotic fluid and the panels were sent either at ambient temperature or at +4°C. In spite of expertise and numerous precautions and controls (including bacteriological controls) in the collection and

preparation of tachyzoites, this procedure was sometimes spoilt with problems, essentially leading to cell death and degradation of the parasite DNA upon arrival at destination. This remained unexplained and made us change to another mode of preparation of the parasites. Since 2008, tachyzoites have been collected from cultivated *T. gondii* (RH strain) maintained in vitro by serial passages in human foreskin fibroblast cell cultures (Besteiro et al., 2008). These have the advantages of being purified, easier to obtain, more standardized, easier to count in a counting chamber and perhaps less subject to degradation than ascites drawn tachyzoites.

These advantages were concomitantly combined with those of freeze-drying (see below).

The choice of the parasite concentrations depends upon the objectives of the EQA and the proficiency of the participants. In France, we opted for low concentrations of parasites (i) because it has been established that a large proportion of infected AFs contain *Toxoplasma* loads < 10 tachyzoites per ml (Costa et al., 2001; Romand et al., 2004), (ii) because diagnostic methods for pathogens are particularly fallible with low concentrations of pathogens in the biological sample (Pelloux et al., 1998), (Chabbert et al., 2004; Chernesky et al., 2002; Kaiser et al., 2007; Lachaud et al., 2001; Lachaud et al., 2002) and (iii) because most participating laboratories were proficient in this diagnosis. The highest concentrations that were ever sent were 100 and 50 tachyzoites per mL. This choice proved right, since these concentrations were always detected by the participants (see below). Conversely, the use of low concentrations come up against the pitfall of the Poisson law of large numbers; this predicts the probability of not having any tachyzoite in the sample tested after aliquoting, as a function of the mean predicted concentration in the initial positive pool. We have experimentally tested this probability by assessing the reproducibility of PCR results among low concentrations samples, which implies that several samples be tested at 'low' concentrations (it is important to keep in mind that the definition of 'low' concentrations is dependent on the sensitivity threshold of the method tested). Such a practice aims at preventing a misevaluation of the method performances. The heterogeneity of distribution which was observed from the PCR results confirmed the Poisson Law, implying that, even if the PCR is realized in triplicate, several samples should be sent to the participants for testing the sensitivity of the methods using concentrations below 5T/mL.

The volume of the sample to be tested should ideally get close to the volume tested in routine practice. However, due to a general lack of matrix fluid, it is difficult to provide each participant with 10-20 mL of fluid, which implies a compromise between the ideal volume and the pragmatic constraints. For this reason, in our EQAs, we have chosen to send 2-mL samples. Others have used 1 mL (Kaiser et al., 2007), which appears even further from clinical practice.

The preparation process of the panels can indeed vary widely between fresh/thawed, frozen or freeze-dried samples. In any case, the samples must be 'as identical as possible' among themselves, which implies several precautions: (i) the parasites must be precisely counted and permanently homogenized during dilution and aliquoting; (ii) the panel samples must be kept at +4°C during the whole process; (iii) all identical samples should be drawn at the same time from the same pools; (iv) after arrival at destination, if in a fluid form, they must be kept at +4°C and processed as soon as possible. It should be noted that no DNA stabiliser or additive (anticoagulants, antibiotics) is used by us at any stage. As said above, freeze-drying offers numerous advantages over the other methods, including robustness, transport at ambient temperature, long-term conservation at +4°C, non-

infectiousness, and flexibility in use, and should be preferred for any EQA using 'fragile' microorganisms. On receipt of the panel, participants store the panel at 2-8°C until processing. Full instructions for reconstitutions of freeze-dried samples are provided. Briefly, lyophilized material should be reconstituted by adding 2 mL of sterile water to each vial. After reconstitution, each cell pellet should be resuspended to the appropriate volume to allow extraction of the complete sample using the DNA extraction protocol developed in the laboratory. After testing the panel, participants may be requested to send their results back and to complete a technical questionnaire.

The composition of the EQA panel must consist of one or several samples containing *T. gondii* and at least one sample negative for *T. gondii*. This minimal requirement has increased over the years of our EQAs. The EQA organized by QCMD sends 10-12 samples using different matrices. Presently, we prefer to send three-five samples containing relatively low concentrations of *T. gondii*. Low concentrations should be sent as several replicate samples. Higher concentrations may be included as positive controls depending on the proficiency of the participants. Negative controls should also be multiplicates, as the carry-over contaminations by amplicons are by nature sporadic.

5. Controls

Due to the importance and multiplicity of samples sent to participants, the panels must be strictly controlled in a number of ways.

PCR-negative amniotic fluid matrixes are pooled and checked again by the routine PCR assay for the absence of *T. gondii* DNA, before being stored at -20°C.

The *T. gondii*-containing panels are checked both for the correct concentrations of tachyzoites and for the absence of DNA degradation. As regards the latter, this can be verified (i) by multiple PCR testing before sending the samples; and (ii) after sending the samples, by keeping an entire panel at room temperature for 5 days in the coordinating laboratory before DNA extraction and PCR testing. However, it is possible that an undetectable level of DNA degradation might contribute to the false negative results observed with very low parasite concentration (e.g. 2-6/mL). In 2006, we had noticed a major unexpected decrease of the *T. gondii* concentration, which made us cancel this edition of the EQA.

Using freeze-drying, the controls are even more stringent, but can be done once and well in advance to the EQA. Pilot experiments were performed to test the influence of freeze-drying and storage on the detectability of *T. gondii* DNA in amniotic fluid. A set of serial dilutions was prepared and aliquoted, along with a negative sample. DNA extraction without freeze-drying (one aliquot of each concentration) was immediately carried out (DNA stored at +4°C) whereas all other aliquots were then freeze-dried. Real-time PCR was simultaneously run in our laboratory on both the non-freeze-dried and the freeze-dried samples (three samples for each). The number of target copies was deduced from the crossing point (Ct) value above which the signal significantly exceeds the baseline. This showed that freeze-drying results in a moderate reduction of Ct values ranging from 0.2 to 1 Ct, whichever the concentrations of tachyzoites tested (from 100.000 to 2 per mL). These differences can then be taken into account in the concentration calculations announced to the participants. Based on these results, three panels are then shipped to four selected reference laboratories. Correct results from all four laboratories for each concentration validate the freeze-drying of the panel.

Freeze-dried samples have been shown to be stable for at least one year storage at +4°C, from PCR testing of panels after 12 months (our unpublished data).

6. Results of the French national EQAs

6.1 Qualitative results

Depending on the year, the whole panel included two to five samples, totaling 765 samples during the 7-year period. Participant results and questionnaires were analyzed on an anonymous basis. The overall results are shown in Table 1.

Date	2002	2003	2004	2005	2007	2008	2009	Total
Number of laboratories	21	22	23	25	28	29	28	176
Number of samples/panel	2	4	5	5	5	5	5	-
Negative samples	21/21	42/44	45/46	50/50	56/56	58/58	56/56	328/331
False-positive^a	0/21	2/44	1/46	0/50	0/56	0/58	0/56	3/331
Positive samples	17/21 (81%)	41/44 (93%)	67/69 (97%)	72/75 (96%)	57/74 (77%)	75/77 (97%)	72/74 (97%)	401/434 (92%)
< 2 T/mL					13/28			13/28 (46%)
2-6 T/mL		19/22	22/23		24/28		54/56	119/129 (92%)
6-10 T/mL	17/21		22/23	23/25		27/29	28/28	117/126 (93%)
20 T/mL		22/22		24/25	26/28			72/75 (96%)
50T/mL			23/23			29/29		52/52 (100%)
100T/mL				25/25		29/29		54/54 (100%)
False-negative^b	4/21	3/44	2/46	3/50	21/56	2/58	2/56	37/434 (8.5%)

Table 1. Overall qualitative results of the French national EQAs for molecular detection of *Toxoplasma* from 2002 to 2009. ^a Number of negative samples found to be positive. ^b Number of positive samples found to be negative

The number of participants in the scheme increased from 21 in 2002 to 28 in 2009 (one centre closed in 2009). All laboratories used an 'in-house' PCR assay with diverse extraction methods. The panel composition and the performance of the participants in relation with the year are shown in Table 1. Data for the entire panel showed a constant progression from the implementation of the national EQA to date. Indeed, correct results were reported in 17/21 (81%) data sets in 2002; 17/22 (77%) in 2003; 20/23 (87%) in 2004; 22/25 (88%) in 2005; 27/29 (93%) in 2008 and 26/28 (93%) in 2009; 2007 was an exception as we wanted to test very low parasite concentrations, which lead to a high number of negative results. The most concentrated positive panels (50 and 100T/mL) were identified correctly in all data sets.

Low positive concentrations (2-6 T/mL) were detected in 86%, 96%, 86% and again 96% of the samples in 2003, 2004, 2007 and 2009, respectively. And very low positives (< 2 T/mL) were detected in 13/28 (46%) of the samples.

The general level of sensitivity in French laboratories was high. The total number of false negative results reported was equivalent to 8.5% of all tests performed on all positive panel members. Except for 2 laboratories, sensitivity problems were observed only for low parasite concentrations (<10 T/ml), where two to four of the 21-29 laboratories reported false negative results. But it should be noted that a single (very low) concentration (< 2 T/ml) was responsible for about half of these false negative results; if one takes out these very low concentrations, false negative results represented 5.4% of 406 samples. Considering the 6-10 T/mL concentration which corresponds to the parasitic load in half of the amniotic fluids according to Costa et al. (Costa et al., 2001), false-negative results represented 7.1% of 126 samples.

It is interesting to note that over the whole period of our study, except in 2007, the number of false negatives decreased over the years from 6.7% in 2002-2004 (Bastien et al., 2007) to 2.7% in 2009. Considering the 6-10 T/mL concentration, the number of false negatives decreased from 19% laboratories in 2002 to 6.9% in 2008 and 0% in 2009 (Table 1). The false-negative results did not appear to be related to a particular PCR method (e.g., real-time, DNA target or primer pair), but can be attributed to technical proficiency, PCR optimization and laboratory practices. Identical conclusions were reached by Kaiser et al. 2007 (Kaiser et al., 2007).

Also noteworthy is the fact that, although we always used lower concentrations than in other EQAs, the percentage of false negative results appeared to be very low as compared with the scarce results published to date. Indeed, a study in 1998 reported 40% of false-negative results at a concentration of 10 T/mL (Pelloux et al., 1998). Another international study involving 33 laboratories (of which less than four seemed to be included in the national network) reported 19.5% false negative results (Kaiser et al., 2007), using concentrations starting from 20 T/mL down to 5 T/mL.

Overall, according to our experience among proficient centers for this molecular diagnosis, and taking into account that all these centers can detect concentrations < 5 T/mL, we consider that a minimum threshold to reach should be defined at 5 tachyzoites per mL based upon a DNA dilution series. This should be confirmed using samples directly extracted at this concentration.

With regard to specificity, the cumulative rate of false-positive results (0.9%; 3/331 samples) was very low in comparison with other molecular EQA programmes, e.g., 2-10% (Kaiser et al., 2007; Walton et al., 2005); 11.6% (Pelloux et al., 1998); or 35% (Valentine-Thon et al., 2001). Indeed, remarkably, only two centers reported a false positive result in 2003; and no false positives have been reported since from any of the centers. Yet, their presence, as well as the fact that the last false positive result observed was from a real-time PCR assay, stress the need to constantly re-evaluate protocols for the prevention of carryover contamination. The multiplicity of replicates of negative controls may help to reveal more contaminations of this type, as they tend to be sporadic in nature.

6.2 Quantitative results

The quantitation of the *T. gondii* amounts in biological fluids is relevant in view of two publications that established a relationship between parasitic loads and either the prognosis or the severity of the infection in the fetus (Costa et al., 2001, Romand et al., 2004). However,

its application in routine practice poses numerous problems and appears as a real challenge to standardization. This is shown on concrete grounds by the fact that, although most proficient centers in France are now equipped with real-time PCR apparatuses (the number of equipped centers rose from 13 to 28 out of 29 between 2004 and 2009), only half in 2008, and a third in 2009, of these centers really quantified *T. gondii* in their results to the EQAs (Sterkers et al. unpublished data).

Relative quantification should be a first step, and should be straightforward as long as the PCR assay shows good repeatability and reproducibility. This may be expressed in number of Cts and can show an increase or a reduction on parasite amounts on a rather fine scale (with an error coefficient corresponding to a delta of Cts ≥ 2). Absolute quantification is more interesting, but poses the problem of the nature and accuracy of the material used for the concentration range. Indeed, (i) ideally, due to different inhibition properties depending on the biological fluid examined, the concentration range testing should be carried out in the same matrix as the one examined (amniotic fluid, blood...); (ii) the nature of the reference material to be quantified in the range may vary widely, from a heterologous plasmid, far from the parasitic target, to plasmids containing inserts of *T. gondii* DNA sequence, and to whole tachyzoites; (iii) the assessment of the amount of targeted DNA before the DNA extraction/PCR should be highly accurate. In any case, it should be kept in mind that the less concentrated is the DNA to be quantified, the less reproducible and accurate is the quantification.

Our experience in the EQAs with respect to quantification reveals wide variations in the results for the same sample. In 2008, for example, quantitative results were reported in 17 (59%) data sets, two of them expressing data as Ct (cycle threshold) values and 15 expressing them as concentrations (tachyzoites/mL). Eleven data sets allowed us calculating the mean, median and standard deviation. Four data sets were excluded as they were considered as off the correct range (variation of at least 10 fold for one point of the range) (Table 2). In 2009, quantitative results were reported in 23 (82%) data sets. But this time, probably due to the communication of the previous year's results and to recommendations from the CNRT, the majority of centers (17) expressed the results as semi-quantitative, whereas 9 laboratories expressed them as concentrations (three laboratories reported both). It is noteworthy that among the four centers which were considered off the correct range in 2008, one (C1) under-evaluated the amounts in 2008 but over-evaluated them in 2009, two did not quantify in 2009, and one did not improve much (Table 2).

The notification of parasitic loads to clinicians remains extremely rare in France, both because of the difficulties encountered in absolute quantification and because there is a general assumption in the community that the two studies quoted above should be confirmed. The need for standardization is here critical. First, there should be a consensus about the matrix and the nature of the DNA target to be used. Second, in order to allow comparison between centers, the estimates of the amounts of targets in the concentration range should be identical for all. Actually, a standardized and accurate concentration range should be distributed to all centers, so that all can work using a robust and reproducible material allowing multicentric comparisons. This is the present project of the CNRT, with the concern of the correlation between the real estimation and the results obtained using serial dilutions of DNA or extractions of diluted suspensions of tachyzoites. Indeed, we repeatedly obtained different results when serially diluting pre-extracted DNA or when directly extracting amniotic fluid spiked with *T. gondii* at different (relatively low) concentrations. A different challenge again will be posed by the implementation of

quantification in whole blood, which might be a useful tool for the follow-up of immunosuppressed patients. Our experience with whole blood shows wide variations in the cycle thresholds, corresponding to variations of parasitemiae by a 1000-fold factor.

		2008			2009		
Real concentration in the samples (tachyzoites/mL)		100	50	10	10	5	5
Centres	C1 ^a	10	4	1	1245	660	142
	C2	20	10	4	15	1	2
	C3	47	20	3	10	5	5
	C4	65	23.7	7.2	23	7	3
	C5	86	30	12	<200 (Ct=30)	<200 (Ct=35)	<200 (Ct=33)
	C6	95	44	10			
	C7	110	50	10	42	7	13
	C8	120 (Ct=29.5)	50 (Ct=30.5)	3 (Ct=34.5)	5	1	3
	C9	126.5	80	15.7	176.5	62	34.5
	C10	135	100	20			
	C11	170	18	2	(Ct=31) ^c	(Ct=36)	(Ct=36)
	C12	225	75	15			
	C13				10	100	10
	C14	350	250	100			
	C15	1100	810	68	100	0	21
	C16	4000	1600	40			
	C17	(Ct=34) ^b	(Ct=36)	(Ct=38)			
	Mean 1 ± SD ^d	444.0±1019.1	211.0±434.4	20.7±28.1	181.2±430.1	61.3±159.4	
	Mean 2 ± SD^e	109.0±57.2	45.5±29.1	9.3±6.0	17.5±13.4	5.2±3.8	
	Median 1 ^d	120	50	10	19	7	
	Median 2^e	110	44	10	12.5	5	
	Max	4000	1600	100	1245	660	
	Min	10	4	1	5	1	

^a Italicized lines and figures show the data which were considered as off the correct range (variation of at least 10 fold for one point of the range)

^b this center did not express the results as an absolute quantification

^c this center did not express the results as an absolute quantification in 2009

^d mean and median calculated with all data sets

^e mean and median calculated without including the data sets considered as "off range": these were found to be very close to the initial amounts estimates

Table 2. Performances of absolute quantification (in tachyzoites/mL) by real-time PCR in French national EQAs in 2008 and 2009

7. Discussion

Our experience with the French national EQAs for the molecular detection of *T. gondii* has been the source of many lessons and improvements. Our researches for an optimal matrix and parasitic material were profitable, and lead to an efficient tool for this unique evaluation of molecular detection for a parasite.

7.1 Diversity of the methods used for molecular diagnosis

The questionnaires associated to the national EQAs for the molecular detection of *T. gondii* have permitted to document the extreme diversity, as well as the evolution, of the methods used, whether for DNA extraction or the PCR itself, as well as of the practices associated to this diagnosis. This has been the subject of a previous article (Sterkers et al., 2009) and another one is in preparation. It is important to stress that the 2002-2009 period of study has witnessed the massive progression of quantitative 'real-time' PCR technology, as opposed to 'conventional' (end-point detection) PCR, as well as, more recently, the emergence of the use of automated DNA extraction. Yet, at the time of writing, all PCR assays used in France are 'in house' (or 'laboratory-developed') assays, and we are therefore far from standardization. As said in the Introduction, this situation, not only are the source of variations in the assays' performances which may be detrimental for patient's care, but also hamper any valuable comparison of data among different centers. Unfortunately, standardization of 'laboratory-developed' methods appears impossible in a world of rapidly changing technologies and wide commercial offer. This will eventually be made possible (i) through a homogenization of pre-analytic and analytic practices, and (ii) through the commercialization of truly highly-performing PCR kits for this pathogen, keeping in mind that these should be evaluated against 'in-house' methods (Morelle et al., submitted). While waiting for the second event, the French CNRT is working towards the first goal.

The French national EQAs have encountered great success and allowed many achievements. They have enhanced communication among proficient laboratories; they have brought the means for authorized centers to self-evaluate their performances; and finally, in part due to the return of information and recommendations attached to the results, they have allowed an improvement in the performances of the PCR assays realized in these centers. The CNRT endeavors to push these improvements further by distributing standard material and establishing recommendations about methods (including DNA primers and targets) and pre-analytic practices.

7.2 Recommendations

The return of information in a yearly report provided the opportunity for writing proposals aiming at improving laboratory performances and practices for this diagnosis. Part of these recommendations have been discussed previously (Sterkers et al., 2009). One of the recent recommendations is to reach a minimal PCR sensitivity threshold, which is being established using a standard biological material. This is not mandatory but, joint to the distribution of benchmark material, should help laboratories to carry out continuous self-evaluation and improve both the performances of their PCR assays and the associated practices. Another important recommendation remains the refined optimization of the PCR conditions (Bastien et al., 2008); indeed, although this appears less necessary with highly-performing modern real-time thermocyclers, we have seen several examples of these where relatively simple changes in the PCR conditions improved the sensitivity of the technique. In

any case, the differences that have been observed among centers in the different EQAs for *T. gondii* appear to be independent of the target, primers or technology, hence to rely essentially on proficiency and care in the optimization of the PCR conditions. Another lesson which has been learnt from these EQAs is the crucial importance of evaluating the diagnostic method as a whole, meaning the DNA extraction together with the PCR method. The former must be adapted to the latter, which means that both should be optimized together (our unpublished data). Thus, even an optimized PCR assay may yield poor results when DNA has been extracted by a suboptimal method for that assay.

7.3 Perspectives

The EQAs for molecular detection of toxoplasmosis hitherto have focused on the diagnosis of congenital toxoplasmosis. In view of the increasing prevalence of toxoplasmosis in immuno-suppressed individuals, there is a clear need to evolve towards the testing of biologically relevant matrices other than amniotic fluid, such as whole blood. It is noteworthy that the molecular diagnosis of CT offers the advantage of being readily confirmed by an independent method (serological follow-up until the age of one year), according to the well defined clinical and biological criteria for CT (Lebech et al., 1996) and updated by (Remington et al., 2004; Tissot Dupont et al., 2003). This makes it possible to calculate sensitivity, specificity, and positive and negative predictive values. Moreover, amniotic fluid is a 'simple' matrix for the PCR, from which spiked mimic samples can easily be made. Conversely, the significance of a positive signal in the blood of immunocompromised patients is much more difficult to analyze, and simulated buffy coat is certainly less easy to produce. Moreover, a blood matrix inevitably yields some degree of inhibition of the PCR, which imposes a supplemental bias in the analysis. It is also important to stress that a PCR assay that has been optimized using amniotic fluid will not necessarily be so when using blood. Finally, as discussed above, another necessary evolution is the assessment of the accuracy of absolute quantification.

In total, EQAs for *Toxoplasma*-PCR have been clearly evolving from a relatively non-professional venture among colleagues to a highly skilled organization. They appear as a necessity for the improvement of molecular diagnosis of infectious diseases, and are becoming compulsory by law in many countries. The progression of molecular diagnosis in Parasitology and Mycology should therefore see a gradual extension of such EQAs.

8. Acknowledgments

We wish to thank Pierre Marty (CHU of Nice), Jacqueline Franck (CHU of Marseille), Hervé Pelloux (CHU of Grenoble), Nathalie Van Langendonck (CHU of Tours) and François Peyron (CHU of Lyon) for their help during the different years of EQAs. We also thank all the centers that participated to these national programs, and more precisely B. Cimon (Angers), P. Millet (Bordeaux), D. Quinio (Brest), M. Vergnaud (Caen), B. Carme and C. Aznar (Cayenne), V. Livrelli and M. Cambon (Clermont-Ferrand), S. Bretagne (Créteil), A. Bonnin and F. Dalle (Dijon), H. Pelloux and M.P. Brenier-Pinchart (Grenoble), L. Delhaes (Lille), M. L. Dardé and D. Ajzenberg (Limoges), F. Peyron (Lyon), J. Franck (Marseille), B. Fortier † (in memoriam) and M. Machouart (Nancy), F. Morio and M. Miegerville (Nantes), H. Yera (Paris-Cochin), P. Thulliez (Paris-Institut de Puériculture), S. Brun and F. Touafek (Paris-Pitié Salpêtrière), P. Roux (Paris-Saint Antoine), J. Menotti and F. Derouin (Paris-Saint Louis), F. Foudrinier and I. Villena (Reims), J.P. Gangneux and F. Robert-Gangneux

(Rennes), P. Flori (Saint-Etienne), D. Filisetti (Strasbourg), M.H. Bessières and S. Cassaing (Toulouse), T.H. Duong (Tours) and O. Eloy (Versailles). We also acknowledge financial support from the French association of hospital practitioners and teachers in Parasitology-Mycolology (ANOFEL) and from the Institut de Veille Sanitaire (through the National Reference Centre for Toxoplasmosis).

9. References

- Al-Soud, W. A. & Radstrom, P. (2001) Purification and characterization of PCR-inhibitory components in blood cells. *J Clin Microbiol*, 39, 485-493.
- Bastien, P. (2002) Molecular diagnosis of toxoplasmosis. *Trans R Soc Trop Med Hyg*, 96 Suppl 1, S205-215.
- Bastien, P., Jumas-Bilak, E., Varlet-Marie, E. & Marty, P. (2007) Three years of multi-laboratory external quality control for the molecular detection of *Toxoplasma gondii* in amniotic fluid in France. *Clin Microbiol Infect*, 13, 430-433.
- Bastien, P., Procop, G. W. & Reischl, U. (2008) Quantitative real-time PCR is not more sensitive than "conventional" PCR. *J Clin Microbiol*, 46, 1897-1900.
- Beld, M. G. H. M., Birch, C., Cane, P. A., Carman, W., Claas, E. C. J., Clewley, J. P., Domingo, E., Druce, J., Escarmis, C., Fouchier, R. A. M., Foulongne, V., Ison, M. G., Jennings, L. C., Kaltenboeck, B., Kay, I. D., Kubista, M., Landt, O., Mackay, I. M., Mackay, J., Niesters, H. G. M., Nissen, M. D., Palladino, S., Papadopoulos, N. G., Petrich, A., Pfaffl, M., W., Rawlinson, W., Reischl, U., Saunders, N. A., Savolainen-Kopra, C., Schildgen, O., Scott, G. M., Segondy, M., Seibl, R., Sloots, T. P., Tang, Y.-W., Tellier, R. & Woo, P. C. Y. (2007) Experts Roundtable: Real-Time PCR and Microbiology. In: *Real-Time PCR in Microbiology: From Diagnosis to Characterization* (ed. Q. A. Ian M. Mackay Sir Albert Sakzewski Virus Research Centre). Caister Academic Press.
- Berger, F., Goulet, G., Le Strat, Y. & Desenclos, J. (2008) Toxoplasmosis in pregnant women in France: trends in seroprevalence and incidence, and associated factors, 1995-2003. *Bulletin Epidémiologique Hebdomadaire*, 14-15, 117-121.
- Besteiro, S., Bertrand-Michel, J., Lebrun, M., Vial, H. & Dubremetz, J. F. (2008) Lipidomic analysis of *Toxoplasma gondii* tachyzoites rhoptries: further insights into the role of cholesterol. *Biochem J*, 415, 87-96.
- Burggraf, S., Reischl, U., Malik, N., Bollwein, M., Naumann, L. & Olgemoller, B. (2005) Comparison of an internally controlled, large-volume LightCycler assay for detection of *Mycobacterium tuberculosis* in clinical samples with the COBAS AMPLICOR assay. *J Clin Microbiol*, 43, 1564-1569.
- Chabbert, E., Lachaud, L., Crobu, L. & Bastien, P. (2004) Comparison of two widely used PCR primer systems for detection of toxoplasma in amniotic fluid, blood, and tissues. *J Clin Microbiol*, 42, 1719-1722.
- Chernesky, M., Smieja, M., Schachter, J., Summersgill, J., Schindler, L., Solomon, N., Campbell, K., Campbell, L., Cappuccio, A., Gaydos, C., Chong, S., Moncada, J., Phillips, J., Jang, D., Wood, B., Petrich, A., Hammerschlag, M., Cerney, M. & Mahony, J. (2002) Comparison of an industry-derived LCx *Chlamydia pneumoniae* PCR research kit to in-house assays performed in five laboratories. *J Clin Microbiol*, 40, 2357-2362.

- Costa, J. M., Ernault, P., Gautier, E. & Bretagne, S. (2001) Prenatal diagnosis of congenital toxoplasmosis by duplex real-time PCR using fluorescence resonance energy transfer hybridization probes. *Prenat Diagn*, 21, 85-88.
- Guy, E. C., Pelloux, H., Lappalainen, M., Aspöck, H., Hassl, A., Melby, K. K., Holberg-Petersen, M., Petersen, E., Simon, J. & Ambroise-Thomas, P. (1996) Interlaboratory comparison of polymerase chain reaction for the detection of *Toxoplasma gondii* DNA added to samples of amniotic fluid. *Eur J Clin Microbiol Infect Dis*, 15, 836-839.
- Kaiser, K., Van Loon, A. M., Pelloux, H., Ferrandiz, J., Picot, S., Wallace, P. & Peyron, F. (2007) Multicenter proficiency study for detection of *Toxoplasma gondii* in amniotic fluid by nucleic acid amplification methods. *Clin Chim Acta*, 375, 99-103.
- King, L., Villena, I., Ancelle, T., Wallon, M., Garcia, P., Thulliez, P. & Goulet, V. (2008) Congenital toxoplasmosis: implementation of a surveillance system in France. *Bulletin Epidémiologique Hebdomadaire*, 14-15, 122-123.
- Lachaud, L., Chabbert, E., Dubessay, P., Reynes, J., Lamothe, J. & Bastien, P. (2001) Comparison of various sample preparation methods for PCR diagnosis of visceral leishmaniasis using peripheral blood. *J Clin Microbiol*, 39, 613-617.
- Lachaud, L., Marchergui-Hammami, S., Chabbert, E., Dereure, J., Dedet, J. & Bastien, P. (2002) Comparison of six PCR methods using peripheral blood for detection of canine visceral leishmaniasis. *J Clin Microbiol*, 40, 210-215.
- Lebech, M., Joynson, D. H., Seitz, H. M., Thulliez, P., Gilbert, R. E., Dutton, G. N., Ovlisen, B. & Petersen, E. (1996) Classification system and case definitions of *Toxoplasma gondii* infection in immunocompetent pregnant women and their congenitally infected offspring. European Research Network on Congenital Toxoplasmosis. *Eur J Clin Microbiol Infect Dis*, 15, 799-805.
- Luft, B. J. & Remington, J. S. (1992) Toxoplasmic encephalitis in AIDS. *Clin Infect Dis*, 15, 211-222.
- Nissapatorn, V. & Abdullah, K. A. (2004) Review on human toxoplasmosis in Malaysia: the past, present and prospective future. *Southeast Asian J Trop Med Public Health*, 35, 24-30.
- Pelloux, H., Guy, E., Angelici, M. C., Aspöck, H., Bessieres, M. H., Blatz, R., Del Pezzo, M., Girault, V., Gratzl, R., Holberg-Petersen, M., Johnson, J., Kruger, D., Lappalainen, M., Naessens, A. & Olsson, M. (1998) A second European collaborative study on polymerase chain reaction for *Toxoplasma gondii*, involving 15 teams. *FEMS Microbiol Lett*, 165, 231-237.
- Remington, J. S. & Desmonts, G. (1995) Toxoplasmosis. In: *Infectious Diseases of the Fetus and newborn infant* (ed. J. S. R. a. J. O. K. Editors). W.B. Saunders, Philadelphia.
- Remington, J. S., Thulliez, P. & Montoya, J. G. (2004) Recent developments for diagnosis of toxoplasmosis. *J Clin Microbiol*, 42, 941-945.
- Romand, S., Chosson, M., Franck, J., Wallon, M., Kieffer, F., Kaiser, K., Dumon, H., Peyron, F., Thulliez, P. & Picot, S. (2004) Usefulness of quantitative polymerase chain reaction in amniotic fluid as early prognostic marker of fetal infection with *Toxoplasma gondii*. *Am J Obstet Gynecol*, 190, 797-802.
- Sterkers, Y., Varlet-Marie, E., Cassaing, S., Brenier-Pinchart, M. P., Brun, S., Dalle, F., Delhaes, L., Filisetti, D., Pelloux, H., Yera, H. & Bastien, P. (2010) Multicentric comparative analytical performance study for molecular detection of low amounts of *Toxoplasma gondii* from simulated specimens. *J Clin Microbiol*, 48, 3216-3222.

- Sterkers, Y., Varlet-Marie, E., Marty, P. & Bastien, P. (2009) Diversity and evolution of methods and practices for the molecular diagnosis of congenital toxoplasmosis in France : a four years survey. *Clin Microbiol Infect.*
- Thulliez, P. (1992) Screening programme for congenital toxoplasmosis in France. *Scand J Infect Dis Suppl*, 84, 43-45.
- Tissot Dupont, D., Fricker-Hidalgo, H., Brenier-Pinchart, M. P., Bost-Bru, C., Ambroise-Thomas, P. & Pelloux, H. (2003) Usefulness of Western blot in serological follow-up of newborns suspected of congenital toxoplasmosis. *Eur J Clin Microbiol Infect Dis*, 22, 122-125.
- Valentine-Thon, E., van Loon, A. M., Schirm, J., Reid, J., Klapper, P. E. & Cleator, G. M. (2001) European proficiency testing program for molecular detection and quantitation of hepatitis B virus DNA. *J Clin Microbiol*, 39, 4407-4412.
- Walton, C., Hawkey, P. M. & James, V. L. (2005) Examination of specimens for mycobacteria in clinical laboratories in 21 countries: a 10-year review of the UK National Quality Assessment Scheme for Mycobacteria Culture. *Clin Microbiol Infect*, 11, 1016-1021.
- Wong, S. Y. & Remington, J. S. (1993) Biology of *Toxoplasma gondii*. *AIDS*, 7, 299-316.