

La PCR quantitative (qPCR) et le guide de bonnes pratiques MIQE : adaptation et pertinence dans le contexte de la biologie clinique

Quantitative PCR (qPCR) and the Guide to good practices MIQE: adapting and relevance in the clinical biology context

Maxime Doods¹
Abalo Chango¹
Afif Abdel-Nour²

¹ Département SNES, Institut Polytechnique LaSalle Beauvais, Unité de recherche EGEAL UP-2012.10.101, France
<abalo.chango@lasalle-beauvais.fr>

² Special Infectious disease unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Résumé. La qPCR a été introduite en recherche clinique et biomédicale comme outil de choix depuis plus de dix ans maintenant. Sa mise en œuvre a été sans cesse croissante dans l'obtention des résultats d'études et de diagnostics. Avec une telle utilisation, la question de la fiabilité et de la pertinence des résultats et informations produits est posée. Cette revue offre une réponse documentée à cette question : le guide de bonnes pratiques MIQE : *minimum information for publication of quantitative real-time PCR experiments*. L'intégralité du processus d'extraction et de l'analyse de l'information biologique y est traitée, avec pour chaque étape une réponse simple. Les contraintes techniques de la recherche ont été intégrées afin de permettre une application plus réaliste de l'application du guide.

Mots clés : *bonne pratique, guide, MIQE, PCR, PCR quantitative*

Abstract. The qPCR has been introduced in clinical and biomedical research for over 10 years from now. Its use in trials and diagnostics is continuously increasing. Due to this heavy use, the question of reliability and relevance of qPCR results has to be asked. This review proposes a documented and evidence based answer to this question, thanks to the MIQE (*minimum information for publication of quantitative real-time PCR experiments*) guideline. The whole analysis process is addressed, from nucleic acids extraction to data management. Simple answers are given, taking into account the technical constraints from clinical research in order to allow a realistic application of this guideline.

Key words: *best practice, guideline, MIQE, PCR, quantitative PCR*

Article reçu le 18 octobre 2013,
accepté le 03 février 2014

La PCR quantitative (qPCR), dont l'ancienne appellation « PCR en temps réel » tend à être abandonnée à cause de la confusion induite par la terminologie anglaise RT-PCR (*reverse transcription PCR* ou *real-time PCR*). La qPCR est un outil puissant pour l'analyse en biologie moléculaire et en biotechnologies. Elle utilise les acides nucléiques (ADN ou ARNm principalement) pour caractériser l'activité d'un système vivant. Elle permet également de quantifier les protéines dans divers échantillons [1].

Son utilisation dans le cadre de la recherche et plus particulièrement en biologie et en clinique est sans cesse croissante

comme le montre l'évolution des publications depuis 2002 (*figure 1*). Ceci s'explique en grande partie par la facilité qu'offre cette technique pour produire des résultats *a priori* pertinents. Cette pertinence perçue est liée à la grande sensibilité de la technique qui permet la détection de variations mineures des séquences d'acides nucléiques (ADN, cDNA, ARN), et également à sa spécificité qui permet de discriminer quantitativement des cibles moléculaires [2]. Actuellement, la PCR digitale en gouttelettes (ddPCR) se développe et se démocratise, offrant un pouvoir discriminant et une sensibilité accrue [3]. Du fait de ces caractéristiques techniques toujours améliorées, des différences infimes au niveau des nucléotides peuvent être mises en évidence. La question de la pertinence biologique de ces résultats est donc d'une importance croissante à mesure que la sensibilité progresse (*figure 1*).

Tirés à part : A. Chango

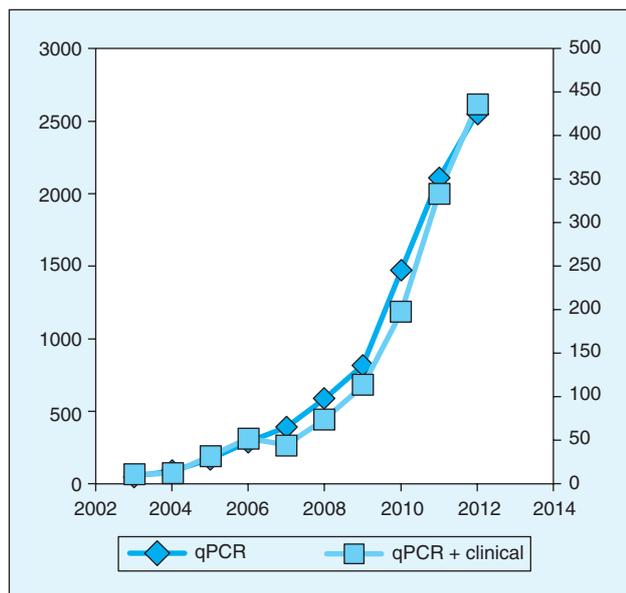


Figure 1. Nombre d'articles publiés par mots clés (qPCR, qPCR + clinique) et par année (source : NCBI).

Dans le cadre de la biologie clinique, il est également important de noter que le recours à la qPCR permet souvent de travailler sur des échantillons dont le prélèvement requiert un désagrément aussi mineur que possible pour le sujet ou le patient. En effet, le volume requis peut être extrêmement faible (< 1 mL), il est parfois possible de travailler avec des *faeces* pour l'étude du microbiote colique [4].

Malgré la performance reconnue de cette technologie, l'utilisation courante ne doit pas occulter la réalité biologique étudiée. S'il est en vérité assez facile de générer des données en qPCR, la garantie de leur qualité et leur adéquation avec la réalité biologique étudiée est autrement plus difficile et pourtant cruciale. Un laboratoire préfère parfois passer sous silence l'absence de contrôle plutôt que de mentionner qu'elle n'a pas effectué ledit contrôle. Ceci nuit à la transparence et laisse la suspicion s'immiscer dans la littérature scientifique.

Récemment, une méthodologie visant à réduire les manipulations nécessaires à une bonne pratique a été publiée [5]. Cette méthodologie recourt à l'estimation de certains paramètres importants pour l'amélioration de la qualité des travaux [6]. Il ne s'agit pas d'imposer quelque norme ni surcharge de travail mais simplement de promouvoir un travail de qualité valorisable à sa juste mesure.

Utilité du guide de bonnes pratiques MIQE

Garantir la qualité des données proposées par la qPCR suppose, d'une part, que l'analyse soit rigoureuse et, d'autre

part, qu'elle soit lisible, reproductible et transparente. Cette problématique est liée à l'utilisation de la qPCR et n'est pas spécifique au contexte clinique. Aussi il semble trivial de se référer à un guide de bonnes pratiques cohérent qui de surcroît existe [7] : le guide de bonnes pratiques *Minimum information for publication of quantitative real-time PCR experiments* (MIQE) (Informations minimales pour la publication d'expériences utilisant la PCR quantitative).

Ce guide publié en 2009 offre une liste d'informations, essentielles ou simplement souhaitables, à porter à la connaissance de l'interlocuteur scientifique. La publication propose une traçabilité de la réalité biologique. Chaque item contribue à garantir non seulement la qualité des résultats, mais aussi leur compréhension et leur adéquation avec le phénomène étudié. Aussi, chaque item est appuyé par une ou plusieurs références qui démontrent la pertinence de chacun. Le guide est souvent synthétisé par un tableau récapitulatif des éléments à vérifier, qui dans un contexte de recherche apporte un côté pratique indéniable. Ce tableau a aussi été adapté dans une application pour tablettes et smartphones, MIQE qPCR (en anglais, disponible gratuitement sur l'app store, Apple). Initialement publié en anglais, ce tableau est maintenant disponible en français (http://www.primersdesign.co.uk/assets/files/miqe_french_poster.pdf).

Présentation du guide

Le MIQE guide propose par essence une ligne de transparence et non une norme. Il est donc légitime de pouvoir en discuter la pertinence et le cas échéant d'y apporter des amendements pour s'adapter à un contexte, clinique par exemple. Sur le point de l'adaptabilité et de la pertinence, il est important de mentionner que le guide ne subit pas d'adaptations par défaut de pertinence. À l'inverse, il tire sa pertinence (au-delà des références bibliographiques sur lesquelles il s'appuie) de son adaptabilité et de sa prise en compte de la complexité intrinsèque des différents contextes expérimentaux. L'essentiel dans ce guide est la volonté profonde d'offrir de la transparence, de la lisibilité et de la pertinence biologique. La liste d'items proposée n'est qu'un socle d'informations contribuant à ces objectifs. Cette liste se présente en différentes sections à détailler.

Nomenclature

Ce point est essentiel dans la mesure où il fixe le vocabulaire employé. L'item « quantification » (et non « quantitation ») ne revêt que peu d'intérêt eu égard à l'absence de cette faute en français (alors qu'il est fréquent en anglais de lire « quantitation » au lieu de « quantification »). En revanche, le remplacement des termes « gène de ménage » et « Ct »

(cycle threshold, cycle seuil) par les termes « Gènes de référence », et « Cq » (Cycle de quantification) revêt un caractère important pour une démarche uniforme.

De même l'appellation qPCR est à favoriser au détriment de RT-PCR. Cette dernière notation induit trop souvent la confusion entre le fait que la PCR soit quantitative (*real time*, RT) et l'étape de rétrotranscription (*reverse transcription*, RT), parfois jusque dans les publications.

Enfin, l'emploi de termes génériques comme « sonde à hydrolyse » en lieu et place de « sondes TaqManTM » permet de s'affranchir de l'utilisation de marques à mauvais escient. Préciser la technologie d'une sonde à fluorescence apporte plus d'information et moins de confusion qu'une marque utilisée comme un générique (surtout quand la marque effectivement utilisée n'est pas celle qui a donné son nom). Eu égard à l'importance que revêt le vocabulaire dans la compréhension d'un texte et à l'information qu'il véhicule, l'ensemble des items de cette catégorie sont essentiels.

L'ensemble de ces termes courants revus par le guide sont rapportés dans la section « nomenclature » du guide MIQE [7].

Plan d'expérience

Dans une expérience – et à plus forte raison dans une étude clinique – le plan d'expérience est une clé de lecture qu'il convient de préciser convenablement pour garantir au lecteur une pleine compréhension du phénomène étudié. La liste n'est encore une fois pas limitative, mais il semble que fournir les effectifs des groupes (en début et fin d'essai le cas échéant) et définir le ou les groupes contrôles en plus des groupes expérimentaux est un minimum indispensable. Par ailleurs, dans une logique de transparence, il est souhaitable de communiquer le réalisateur effectif (interne ou prestataire extérieur) des différentes étapes d'un essai et la contribution effective des différents auteurs. Dans un contexte de la biologie clinique, il est évident que ces informations initiales gagneront à être complétées par d'autres informations complémentaires : intervention ou observation, en parallèle ou *cross over*, simple ou double aveugle etc.

Echantillonnage

Le contexte linguistique et expérimental étant posé, il convient de préciser la nature et les traitements subis par les échantillons analysés. Ceci est primordial dans la mesure où leur manipulation peut induire leur altération et potentiellement briser la chaîne de l'information biologique entre l'échantillon au moment du prélèvement et le résultat obtenu en qPCR. Parmi les informations à fournir on citera la nature de l'échantillon, la procédure d'échantillonnage ayant conduit à l'obtention des échantillons, leur éventuel

protocole de fixation (fixation sur paraffine, surgélation), ou leurs conditions de stockage (durée, température, si pertinent conditions physico-chimiques). La majeure partie des informations de cette catégorie sont essentielles. Seul le volume ou la masse d'échantillon utilisé est souhaitable et non essentiel.

Extraction des acides nucléiques

Dans l'hypothèse où l'information soit arrivée intacte à ce stade de l'expérience, il est crucial de ne pas la dégrader lors de son extraction et de l'extraire pleinement afin de ne pas commettre si l'on peut dire un mensonge par omission. Aussi de nombreux détails sont pertinents quant à cette étape. Préciser que « les ARNm ont été extraits conformément aux instructions du fabricant » ou « en suivant les préconisations de Dooms *et al.*, 2013 » ne suffit pas. Une publication est faite pour communiquer de l'information scientifique, pas pour faire un jeu de piste. Un protocole d'extraction est parfois complexe, offre plusieurs possibilités. Il arrive également que la publication à laquelle renvoie l'article contienne elle-même un renvoi, voire qu'*in fine* l'information ne soit pas suffisante. Pour peu que le lecteur ne soit pas abonné à la bonne revue ou que sa patience montre rapidement ses limites, l'information s'arrête aux portes du processus de diffusion.

Aussi, il est essentiel de centraliser les informations dans un rapport ou un tableau lié à votre publication. Concernant les informations pour l'extraction d'acides nucléiques figurent celles du kit utilisé et/ou l'instrumentation associée à l'extraction, les éventuelles modifications apportées (en détails, pas seulement « kit X modifié » qui n'apporte aucune précision si ce n'est que l'information est perdue), l'origine des réactifs supplémentaires le cas échéant, les traitements de type DNases ou RNases, la vérification d'absence de contamination (ADN ou ARN), la qualité de l'acide extrait (intégrité, pureté), la vérification d'absence d'inhibiteurs (par des séries de dilutions par exemple).

Rétrotranscription

Dans l'hypothèse d'un travail fondé sur l'analyse d'ARN, il est dans l'immense majorité des cas nécessaire d'effectuer une rétrotranscription (que l'on notera sans vergogne « RT » puisqu'ayant mentionné la PCR quantitative avec l'appellation qPCR et non RT-PCR). Là encore, la traçabilité de l'information biologique est essentielle. Elle est de surcroît assez facile à assurer. Il suffit de mentionner la quantité d'ARN utilisée pour la rétrotranscription, l'enzyme et sa concentration (ou simplement le kit utilisé), le type d'amorces utilisées (oligo dT, amorces aléatoires), les conditions (temps, températures, volume réactionnel). Bien que ce ne soit pas essentiel, la communication des Cq

avant et après RT et les conditions de stockage des ADNc est souhaitée.

Cible de qPCR

Arrivé à ce stade, pour peu que les informations mentionnées ci-dessus aient été communiquées, le lecteur est en mesure d'apprécier la conservation de l'information. Reste à lire cette information. Là encore la transparence est de mise. L'idéal est bien entendu la présence de l'accession number du gène ciblé, avec la mention du ou des variants d'épissage le cas échéant et la situation de la séquence cible (en particulier l'emplacement vis-à-vis des séquences introniques ou exoniques). La longueur et les structures secondaires pouvant influencer l'efficacité de l'amplification, ces informations sont également à communiquer.

Au-delà des aspects techniques liés à la cible à amplifier, il est important de s'assurer qu'elle est bien spécifique, qu'elle ne comporte pas dans le génome ou le transcriptome d'homologues (pseudogènes, rétro-pseudogènes). Un mauvais ciblage produirait en effet un résultat parfaitement lisible, mais sans lien suffisamment étroit avec le phénomène étudié, ou pire contradictoire avec la réalité biologique.

Oligonucléotides

Ce sont eux qui vont permettre de lire l'information. Ils sont le bras moteur de la partie précédente (cible de qPCR). De la même façon qu'il est essentiel d'avoir une pharmacie efficace pour mettre en branle les traitements proposés par la médecine, il est important de s'assurer de la qualité et de la pertinence des oligonucléotides pour une cible donnée. Le paramètre le plus essentiel et pourtant pas le plus fréquent est la publication des séquences utilisées. Lors de l'utilisation de kits commerciaux les références de ces derniers doivent être fournies, en évitant dans la mesure du possible les boîtes noires qui proposent des essais prêts à l'emploi sans précisions sur les séquences ou au moins leur contexte. Lors de l'utilisation d'amorces ou de sondes modifiées, il est crucial de communiquer précisément les modifications effectuées. Enfin, dans un but de lisibilité de l'expérience et de reproductibilité, les informations sur le fournisseur et la méthode de purification des amorces/sondes doivent être transmises.

Protocole de qPCR

Pour filer la métaphore médico-pharmaceutique, nous supposons ici être en présence d'un bon diagnostic (contexte expérimental), d'une bonne prescription (cible de qPCR) et d'un traitement adapté (oligonucléotides). Reste à préciser le mode d'administration du traitement, ici le protocole

de qPCR. Tous les éléments étant réunis, reste à les faire réagir de façon adéquate. Les éléments essentiels sont la nature et la concentration de l'enzyme utilisée, la quantité d'ADN ou ADNc utilisée pour chaque réaction, la composition du milieu réactionnel (tampon, concentration en Mg, dNTPs, éventuels additifs) et les conditions complètes de réaction (machine utilisée, volume et programme de thermocyclage). Par ailleurs, en ce qui concerne la qualité des plastiques pouvant influencer la réaction, il est bon de la préciser. Enfin, dans un contexte clinique où la quantité d'échantillons à manipuler est souvent importante, le fait que la mise en place des réactions soit manuelle ou robotisée est important. On peut en effet concevoir que l'assistance d'un automate permette de réduire les erreurs de pipetage et ainsi d'augmenter la répétabilité des résultats. Cet effet croît avec la taille de la population analysée.

Validation de la qPCR

À ce stade, l'information brute est livrée. Il est tentant de passer directement à l'analyse pour obtenir au plus tôt ce que l'on pense être un résultat. Néanmoins, quelle serait la valeur d'un résultat issu de données brutes non validées ? Nulle.

En complément – et non en remplacement ! – des informations mentionnées ci-avant, des vérifications simples permettent de valider les données brutes. La qPCR est-elle bien spécifique (vérifiable avec une courbe de fusion ou une électrophorèse) ? La fluorescence détectée lors de l'utilisation de SYBR Green est-elle liée à la création d'un produit de PCR (quelles sont les valeurs de C_q pour les échantillons témoins) ? La qPCR est-elle efficace et fiable tout au long de la gamme de valeurs observées (quelles sont pour chaque essai les variations au long de la gamme d'efficacité, particulièrement pour les valeurs basses, quelle est l'efficacité lisible sur la courbe et quel est le R^2) ?

Répondre à ces questions ne prend que peu de temps. Les réponses dépendent globalement de la nature de l'échantillon et de l'essai, et ne varient que peu entre plusieurs individus.

Analyse des données

Enfin, le temps de la valorisation des données brutes (validées) en résultats exploitables est venu. Comme tout au long de l'expérimentation, la transparence va permettre aux lecteurs d'apprécier la qualité de l'expérimentation. Les points à mentionner comportent le programme utilisé pour l'analyse (et sa version !), la méthode de détermination du C_q , le nombre de gènes de référence utilisés et la justification de ce choix (on trouve encore trop souvent des données normalisées avec un seul gène pris par usage et non validé dans le contexte étudié), la méthode utilisée pour normaliser les données et les dispositions prises dans le cas de valeurs

suspectes. Concernant l'analyse, elle ne sera complète que si la répétabilité et la reproductibilité de l'expérience sont estimées (stade et nombre des réplicats techniques et biologiques), et si les paramètres statistiques essentiels indiqués (calcul de puissance, tests statistiques utilisés). La perfection serait approchée si les données brutes étaient soumises au format RDML, prévu pour la communication de ce type d'information.

Impact sur la recherche clinique

Dans la mesure où l'ensemble des informations mentionnées ci-dessus tend à être communiquées, l'intégrité et la pertinence des données doivent être garanties. Le contexte clinique impose par nature une exigence dans la qualité de la conduite expérimentale et la garantie du suivi des données. Dans cette optique, certains points de contrôle requièrent des manipulations supplémentaires, ce qui dans un contexte de tension des ressources peut ralentir l'adoption de ce guide. La transparence ne requiert en théorie pas d'expériences supplémentaires. Cette méthodologie recourt à des estimations de certains paramètres, réduisant le coût de la qualité. Elle permet d'ajuster le degré de vérification de la qualité selon les ressources disponibles et ainsi en toute transparence de communiquer sur les efforts entrepris pour garantir la traçabilité des résultats. L'objectif est de laisser place à la curiosité en lieu et place de la suspicion. Ceci favorise la critique constructive et permet en s'assurant de l'adéquation entre le phénomène biologique étudié et la donnée générée de valoriser la recherche scientifique en connaissance de cause.

Conclusion

Dans un contexte d'utilisation courante de la qPCR, il apparaît nécessaire de proposer un guide cohérent de bonnes pratiques. Ceci est d'autant plus nécessaire que de nouvelles techniques fondées sur la qPCR voient le jour. On citera particulièrement la ddPCR qui apporte une sensibilité accrue. Cette sensibilité toujours en hausse et dont on peut *a priori* se réjouir, représente néanmoins un risque si l'information n'est pas tracée biologiquement. La sur-interprétation de

résultats faussement intéressants peut induire en erreur. La conciliation de l'exigence de résultats de la recherche biomédicale face aux enjeux de santé publique et à celle non moins impérieuse de se cantonner aux ressources disponibles semble intrinsèquement impossible. En effet, la rigueur scientifique nécessaire à la parution de résultats valorisables en termes de santé publique est consommatrice de ressources (taille des échantillons, multiples réplicats). Le guide MIQE met en lumière les points essentiels à une bonne lisibilité des articles scientifiques, et la méthodologie évoquée permet d'estimer ces informations sous forte contrainte de ressources. Nous espérons que ce guide dont l'application récente commence par s'imposer progressivement apportera une grande souplesse et garantie dans le croisement des résultats et dans le cas des études multicentriques.

Liens d'intérêts : Les auteurs déclarent ne pas avoir de lien d'intérêt en rapport avec cet article.

Références

1. Abdel Nour AM, Aussenac T. Current advances in proteomics. *Pharma Focus Asia* 2008 ; 7 : 24-5.
2. Chango A, Nour AA, Bousserouel S, Eveillard D, Anton PM, Gueant JL. Time course gene expression in the one-carbon metabolism network using HepG2 cell line grown in folate-deficient medium. *J Nutr Biochem* 2009 ; 20 : 312-20.
3. Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, Haynes R, *et al.* The digital MIQE guidelines : minimum information for publication of quantitative digital PCR experiments. *Clin Chem* 2013 ; 59 : 892-902.
4. Lecerf JM, Depeint F, Clerc E, Dugenet Y, Niamba CN, Rhazi L, *et al.* Xylo-oligosaccharide (XOS) in combination with inulin modulates both the intestinal environment and immune status in healthy subjects, while XOS alone only shows prebiotic properties. *Br J Nutr* 2012 ; 108 : 1847-58.
5. Dooms M, Chango A, Barbour E, Pouillart P, Abdel Nour AM. Improving biological relevancy of transcriptional biomarkers experiments by applying the MIQE guidelines to pre-clinical and clinical trials. *Methods* 2013 ; 59 : 147-53.
6. Bohlenius H, Eriksson S, Parcy F, Nilsson O. Retraction. *Science* 2007 ; 316 : 367.
7. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, *et al.* The MIQE guidelines : minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 2009 ; 55 : 611-22.