Short Communication

Improving the standardization of mRNA measurement by RT-qPCR

Rebecca Sandersa,*, Stephen Bustinb, Jim Huggeta, Deborah Masonc

a Molecular and Cell Biology, Science and Innovation, LGC, Queens Road, Teddington, Middlesex TW11 0LY, UK
b Faculty of Medical Science, Anglia Ruskin University, Michael Salmon Building, Chelmsford, Essex CM1 1SQ, UK
c School of Biosciences, Cardiff University, Sir Martin Evans Building, Museum Avenue, Cardiff CF10 3AX, UK

A R T I C L E   I N F O
Handled by Justin O’Grady
Keywords: Standardization RNA RT-qPCR Reference materials Normalization Error

A B S T R A C T
Human health and safety depend on reliable measurements in medical diagnosis and on tests that support the selection and evaluation of therapeutic intervention and newly discovered molecular biomarkers must pass a rigorous evaluation process if they are to be of benefit to patients. Measurement standardization helps to maximize data quality and confidence and ultimately improves the reproducibility of published research. Failure to consider how a given experiment may be standardized can be costly, both financially as well as in time and failure to perform and report pre-clinical research in an appropriately rigorous manner will hinder the development of diagnostic methods. Hence standardization is a crucial step in maintaining the integrity of scientific studies and is a key feature of robust investigation.

Our recent review of the literature [18] has shown that the qPCR data underlying the vast majority of publications reporting use of this technique are, at the very least, inadequately reported and that the peer review process allows the publication of incomplete experimental protocols, yielding results that are difficult to evaluate independently. An analysis of all colorectal cancer publications that made use of qPCR between 2006 and 2013 shows that only 3% (n = 179) report sufficient experimental detail to allow a reliable assessment of the qPCR data [19]. That paper also showed that 92% of publications used a single reference gene, with 13% validating its use and 92% of papers use a method of analysis that is meaningless unless PCR efficiency is known, yet 82% do not mention PCR efficiency. A more recent analysis found that 95% of papers (n = 20) used a single reference gene for normalisation, with only 20% using a single validated reference gene [20]. Two other surveys found that 100% of papers (n = 20) used inappropriate analysis and normalisation procedures [21,22]. Other errors, such as incorrect use of controls or even wrong selection of primer sequences, may not be identified by the peer review process. Re-assessment of some of these studies has led to publication retractions [23–26], encouraged misleading conclusions and wasted both time and money of funders and researchers following up on these data [18]. Such errors may be particularly difficult to identify when applying findings from one model system to another, for example, findings from cell line based studies or animal models being compared and/or applied to human clinical samples. Discardance in clinical findings compared to a model system may be put down to failings in the model system, when in fact, the study protocols may be at fault.

* Corresponding author.
E-mail address: rebecca.sanders@lgcgroup.com (R. Sanders).
Retractions, however, are rare and evidence suggests that this may be because the original data have not been questioned/re-evaluated rather than because such mistakes are infrequent [18]. When comprehensive reporting practices are not followed, it becomes difficult not only for such results to be successfully reproduced, but also to determine where the causes of error lie when they arise. Indeed, it may be the case that a second, conflicting study has difficulty getting published at all. Whilst it is accepted that authors of conflicting investigations should collaborate in order to resolve their differences, this process would be greatly simplified if all researchers followed recommended practices and validated procedures within their own laboratories, as suggested in this manuscript, and complied with accepted guidelines (such as MIQE [6]) when reporting data. Of course, it is also possible that some conflicting reports represent the normal biological variation, but in this instance the significance of the different findings will have a direct impact on what one can conclude from either study, and further work would be required to understand such discrepancy.

The sources and degree of error should be described and, where necessary reduced, when validated normalization and best practice procedures are in place. Our studies employing dPCR have highlighted discrepancies that cannot always be detected by qPCR, but may still influence a measurement result. For example, we have previously demonstrated that dPCR is capable of making precise measurements of synthetic and endogenous RNA molecules in a complex RNA background [27]. A possible underestimation bias exists for RNA measurements, with RT-dPCR quantification being significantly lower than that derived from UV spectrophotometry [27]. RT-dPCR was also shown to be more precise than reverse transcription (RT)-qPCR, with RT-dPCR able to highlight the template and assay-specific biases of different one-step RT-qPCR kits. This demonstrated that cDNA prepared using different RT enzymes, or primed with different primers, may result in different yields. For accurate analysis using different approaches, the inherent biases must be understood and factored into a comparison between different PCR approaches or other methods, such as UV spectrophotometry, capillary electrophoresis or fluorescence-based measurement.

Different extraction methods recover RNA with differing efficiencies, with matrix effects, cell debris and/or the quantity of input material, such as the number of cells, having an influence [28–33]. Extraction efficiency is particularly relevant when evaluating transcripts of low abundance and so experiment-specific validation is important. Furthermore, kit-dependent co-purified contaminants and extraction buffer compounds can affect downstream applications such as efficiency of DNase treatment and RT-qPCR. Both extraction kit/chemistry and yield determination must be assured in a method-specific manner for sample sets to remain equally valid.

When there is maintenance of experimental consistency and an understanding of sources of bias and error, theoretically any methods may be expected to generate reproducible results. To assist in this endeavour we provide a best practice guide to aid in improving the assessment of sources of error and the translation to reducing experimental error.

Based on the findings in our previous studies [27,34–36], the authors make the following recommendations for the evaluation of mRNA levels by RT-qPCR:

1. **PCR is arguably the most sensitive molecular method available today as it has the potential to detect single molecules; consequently PCR is susceptible to low level contamination. Furthermore, as a PCR generates billions of copies of the very molecule of interest, this poses an additional significant contamination risk. High copy number plasmids and synthetic templates should also be considered a contamination risk on the same level as PCR product. Consequently, laboratories should comprise areas that separate the experimental setup (pre-PCR) from subsequent analysis (post-PCR).**

Separation of RNA/DNA extraction from PCR reagent preparation (pre-sample mix) will further reduce contamination risk.

- It is equally important to change lab coats between these areas to reduce contamination.
- When handling RNA samples at any stage prior to qPCR, it is advisable to employ experimental procedures such as routine treatment of surfaces/racks/pipettes etc. with a solution to remove RNases (such as RNaseZap, Ambion, or similar alternatives).
- The use of RNase-free plasticware and water is essential.

mRNA measurement by RT-qPCR

- For accurate quantification of a given transcript, replication should be performed at stages that reflect the largest source of error, where this is practical. For example, where an experiment compares the effect of treatment on a cell culture ideally the experiment should replicate the culture flasks. Where clinical samples have been processed and are available as RNA extracts, it is prevalent to convert the RNA to cDNA and rely on this sample for the remainder of the study. However, as the RT step can provide considerable variance, reliance on a single reverse transcription experiment may lead to bias generating results that are challenging to reproduce. Working with and replicating the RNA sample, and therefore the RT, can lead to a more accurate estimation of the associated error than when just replicating the qPCR on a single cDNA sample.

- If replication can only be performed at limited stages due to cost/resources/tissue availability, it should be focussed at earlier steps in the protocol, i.e. subject/independent experiment replicates should hold the highest priority.

- Samples may be stored at −80 °C as either lysate or extracted total RNA. For clinical samples, material should be snap frozen in liquid nitrogen and stored at −80 °C until processing. Storage at −80 °C may not be required if a suitable preservative solution, such as RNAlater or equivalent, is applied.

- For total RNA extraction, different methods can yield different amounts of material and matrix components. As such, extraction methods should be validated on an experimental-specific basis.

- To reduce the effect of repeated freeze/thaw cycles, total RNA extracts should have a practical number of aliquots prepared and stored at −80 °C.

- To protect total RNA from degradation, it is advisable for it to be diluted in a stabilising agent such as RNA storage solution (Ambion), or equivalent. This can be a problem with any sample source, but clinical samples are of special concern because of their complexity and potential inconsistencies in sample size, collection, storage and transport can lead to variable quality of RNA templates [34,37,38].

- Total RNA may need to be DNase treated. Where this is the case it is recommended to assess for DNase efficiency, for example by using Alu PCR [39,40] or equivalent (samples pre and post DNase treatment).

- mRNA should be tested for integrity, using methods such as a 5′-3′ assay, for a highly abundant target [41].

- Extracts should be assessed for co-extracted inhibitors, which can be tested using the SPUD assay [42,43], or equivalent. However, it is important to remember that different PCR reactions can be differentially affected by inhibitors [44].

- RT may be performed using either one-step or two-step processes. When measuring low abundance targets, oligo d(T)16 or gene specific RT priming is recommended (as opposed to random primers), based on specificity for mRNA or specified targets, respectively. However, when using difficult samples such as formalin-fixed paraffin-embedded (FFPE) material, the best approach should be determined on an individual experiment basis.

- Carrier RNA (such as Yeast total RNA) can be added pre-extraction to aid in the recovery of low abundance targets or targets from minimal starting material.
Carrier RNA can be added to the RT reaction (both one-step and two-step processes) to aid linear performance of the RTase.

The validity of the RTase should be tested for each experimental purpose, particularly when evaluating low abundance targets. When measuring low abundance mRNA species, one-step RT-qPCR and RTases lacking RNase H activity should be considered (and validated).

If a calibration curve will be used for quantification of a target, reference sample dilutions should be performed at the RNA (not cDNA) stage (at a minimum n = 3 replicates per dilution point). Furthermore, template type should be maintained where possible, for example, linearized plasmid should be quantified using a standard curve prepared from linearized plasmid.

External RNA controls, for measurement of a calibration curve or normalisation purposes, should be spiked into matrix-matched samples.

Both positive and negative controls should be utilized throughout the experimental process.

qPCR is highly robust and precise and, once optimised, does not necessarily need to be replicated. However, if a new batch of primers or probes or a new lot of reagents is introduced, it is advisable to compare the performance of the assays to ensure comparable results are achieved.

**qPCR Assay design:**

- If DNase treatment is performed and validated by Alu PCR, assays do not necessarily need to be designed to cross an exon–exon boundary, although this approach is still preferred where possible. DNase treatment is particularly important when assays cannot be designed to cross and exon–exon boundary, for example, in single exon genes or pseudogenes. RT negative controls should always be employed as standard to identify any contaminating DNA.
- All assays should be evaluated for specificity. Where possible when performing qPCR, probes such as hydrolysis probes, molecular beacons and scorpion probes, should be used to ensure additional specificity. Where intercalating dyes are used, melt curve analysis should always be performed to evaluate specificity.
- An evaluation of PCR efficiency, whether by standard curve or any other of the available methods is essential. Assay efficiencies may also be estimated using amplification curve fitting algorithms, which are dependent on the number of cycles over which there is an increase in fluorescence, and several such approaches have been proposed [45–49]. Any such approach must first be validated.

**Additional considerations:**

- For analysis of cell line gene expression, cells of equivalent age should be evaluated. At a minimum, comparable cell passage, but ideally cumulative population doubling of cells should be matched throughout a study.¹
- All possible aspects of an experimental set-up should be controlled, e.g. consistency in sample source/type-processing/storage, reagent batches, instrument calibration etc.
- Meaningful comparisons can only be made where the same experimental set-up, reagents and methods are used.
- Published data must include all sample and experimental details to facilitate data reproduction and comparison, either in the main text or in supplementary files.
- All sources of variability should be considered and accounted for before conclusions are made. Especially important is discriminatory power of the particular experimental process, i.e. the ability of a particular experimental design to distinguish between different outcomes. If the technical variability of a particular measurement result is experimentally determined to be 5%, then a biological difference of 2% cannot be determined by this experimental set-up. An estimate of measurement uncertainty should be used to convey the confidence in a measurement result [50].

These recommendations are summarised in Supplementary Table 1. Measurement uncertainty has two components: systematic and random variation [50]. Systematic errors lead to bias in the measurement. These error components are fixed and predictable and may be inherent to various instruments and methods. Random variation occurs when making repeated measurements. This is related to precision; a measure of the degree of agreement between replicate measurement results obtained for the same sample. Contributing factors are multiple and include issues of sampling, different analysts as well as each stage of the stepwise protocol necessary for a measurement [50]. Including an estimate of measurement uncertainty when reporting values allows comparisons to be made between samples and between data sets, where appropriate.

Where available, reference materials (including, but not limited to Certified Reference Materials) can be used to assist reproducibility between laboratories, allowing for harmonization of data. dPCR could assist in the improvement of qPCR through accurate value assignment of reference materials, and to determine RT enzyme performance [27] and ensure that the RT enzyme generates the same result as the PCR enzyme, i.e. counting RNA molecules per cell without the introduction of bias. This approach may also be applied by reagent manufacturers to examine extraction procedures. The absolute counting possible with dPCR methods will allow simple determination of when an RT enzyme or RNA extraction method is not working efficiently and so preventing biases which cannot always be defined using qPCR. It also offers a higher level of precision, when such measurements are needed to determine small fold-changes. Such scrutiny of measurement processes will enable identification of biases and so result in improvements to the experimental protocol. Application of dPCR in this way, as a kind of process evaluator, would help to keep those processes faithful to the measurement. It would also allow an improvement of all process steps and aid harmonisation of protocols and ultimately improve inter-laboratory reproducibility. It is important to note that while reproducibility may be improved, this does not necessarily indicate that a result is accurate.

Important continued efforts include the education and uptake of standardized laboratory and reporting practices as well as adherence to recommended guidelines when reviewing manuscripts and grant applications. When scientists have been doing things a particular way for years it can be difficult to convince them to change. Seasoned researchers may be resistant to change, especially with funding pressures and demands to publish. Nevertheless, both compliance with, and dissemination of guidelines could be enforced by the requirement of such standards in manuscripts by publishing journals and funding bodies. This approach has been successfully employed for animal experiments (ARRIVE guidelines) [51] and microCT imaging/reporting (ASBMR guidelines) [52]. Furthermore, capturing training and newly qualified young scientists, and ensuring they are schooled in the importance of such standardization approaches, will go a long way to safeguarding the integrity of future scientific endeavor.

The work referenced here further highlights the need for standardization in all aspects of methodology and may be used to guide development of studies investigating sources of variability. The employment of dPCR value assigned calibrant materials (reference samples) would facilitate greater accuracy for absolute quantification by qPCR. Efforts should focus on internal target standardization approaches, whether that is using multiple reference genes, expressed repetitive elements (such as Alu repeats [40]), an alternative approach or a

¹ Cumulative population doubling: total number of population doublings of a cell line since their primary isolation in vitro. An intrinsic measure of the “age” of the particular culture.
behavioural plasticity in the Australian plague locust, BMC Mol. Biol. 12 (2011) 7.
[13] T. Hruz, M. Wys, M. Doquier, M.W. Pfaffl, S. Masanetz, L. Borghi, P. Verbrugghe, L. Kalaydjieva, S. Bleider, O. Laule, P. Descamps, W. Gruissem, P. Zimmermann, RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalisation, BMC Genom. 12 (2011) 15.
[14] O. Fedrigo, L.R. Warner, A.D. Pfefferle, C.B. Babbitt, P. Cruz-Gordillo, G.A. Wray. A pipeline to determine RT-qPCR control genes for evolutionary studies: application to primate gene expression across multiple tissues, PLoS One 5 (9) (2010).
[15] M.B. Maess, S. Sendelbach, S. Lorowski, Selection of reliable reference genes during THP-1 monocyte differentiation into macrophages, BMC Mol. Biol. 11 (2010) 90.
[16] B. Zyzynska-Granica, K. Kozlak, Identification of suitable reference genes for real-time PCR analysis of statin-treated human umbilical vein endothelial cells, PLoS One 7 (10) (2012) e51547.
[17] A. Arun, V. Baumle, G. Amelot, C.M. Nieberding. Selection and validation of reference genes for qRT-PCR expression analysis of candidate genes involved in cell communication in the butterfly bicyclus anynana, PLoS One 10 (3) (2015) e0120401.
[18] S.A. Bustin, V. Benes, J. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G. Shipley, C.T. Wittwer, F. Schiering, P.J. Day, M. Abreu, B. Agudo, J.F. Beauclair, A. Beekers, S. Bogaert, J.A. Browne, F. Carrascos-Ramiro, L. Ceelen, K. Giborowski, P. Cornil, S. Coulon, A. Cuypers, S. Des Selves, J. De Cuninck, J. De Cuyper, H. De Spieghelere, R. Deckers, A. Dhedene, K. Durinc, M. Ferreira-Texeira, A. Fleur, J.M. Gallup, S. Gonzalo-Flores, K. Goossens, F. Heindryckx, E. Herring, H. Hoiland, L. Icardi, R. Jaggi, F. Javad, M. Karampelias, F. Kibenge, M. Kibenge, C. Kumps, I. Lambertz, T. Lammens, A. Marky, P. Messiaen, E. Mets, S. Morais, A. Murdasa-Rubro, J. Nakwila, H. Nelis, P.A. Olovic, K. Ren, M. Plungis, M. Ribana, P. Rodriguez-Santos, P. Roudou, R. Sanders, K. Schmed-Biekle, K. Skovgaard, K. Smeets, L. Tabera, S. Toegel, T. Van Acker, V. Van den Broeck, J. Van de Meulen, M. Van Gele, G. Van Peer, M. Van Poucke, N. Van Roy, S. Vergult, J. Wauman, M. Tsibinuka-Wikander, E. Willems, S. Zaccara, F. Zeka, J. Vandesompele, The need for transparency and good practices in the qRT-PCR literature, Nat. Methods 10 (11) (2013) 1063–1067.
[19] J.R. Dijkgraaf, L.C. van Kempen, I.D. Nagtegaal, S.A. Bustin, Critical appraisal of quantitative PCR results: can we rely on published qPCR results? Mol. Oncol. 8 (4) (2014) 813–818.
[20] S. Bustin, T. Nolan, Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research, Eur. J. Clin. Invest. 47 (10) (2017) 756–774.
[21] S.A. Bustin, The reproducibility of biomedical research: sleepers awake!, Biomol. Detect. Quantif. 2 (2015) 34–42.
[22] S.A. Bustin, T. Nolan, Improving the reliability of peer-reviewed publications: we are all in it together, Biomol. Detect. Quantif. 7 (2016) A1–5.
[23] H. Bohlenius, S. Eriksson, F. Parcy, O. Nilsson, Retraction, Science 316 (5823) (2007) 367.
[24] R.N. Kanadia, C.L. Cepko, Retraction Alternative splicing produces high levels of noncoding isoforms of BHLH transcription factors during development, Genes Dev. 16 (9) (2002) 1145–1149.
[25] N. Ikeda, Retraction of Enhanced expression of coproporphyrinogen oxidase in malignant brain tumors: a genetic paradigm for the lack of reproducibility in molecular research, Eur. J. Clin. Invest. 33 (92) (2003) 571–572.
[26] J.R. Dijkstra, L.C. van Kempen, I.D. Nagtegaal, S.A. Bustin, Critical appraisal of quantitative PCR results: can we rely on published qPCR results? Mol. Oncol. 8 (4) (2014) 813–818.
[27] S. Bustin, T. Nolan, Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research, Eur. J. Clin. Invest. 47 (10) (2017) 756–774.
[28] S.A. Bustin, The reproducibility of biomedical research: sleepers awake!, Biomol. Detect. Quantif. 2 (2015) 34–42.
[29] S.A. Bustin, T. Nolan, Improving the reliability of peer-reviewed publications: we are all in it together, Biomol. Detect. Quantif. 7 (2016) A1–5.
[30] H. Bohlenius, S. Eriksson, F. Parcy, O. Nilsson, Retraction, Science 316 (5823) (2007) 367.
[31] J.R. Dijkgraaf, L.C. van Kempen, I.D. Nagtegaal, S.A. Bustin, Critical appraisal of quantitative PCR results: can we rely on published qPCR results? Mol. Oncol. 8 (4) (2014) 813–818.
[32] S. Bustin, T. Nolan, Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research, Eur. J. Clin. Invest. 47 (10) (2017) 756–774.
[37] C.A. Perez-NoVo, C. Claeys, F. Speleman, P. Van Cauwenberge, C. Bachert, J. Vandesompele, Impact of RNA quality on reference gene expression stability, Biotechniques 39 (1) (2005) (52, 54, 56).
[38] M.A. Valasek, J.J. Repa, The power of real-time PCR, Adv. Physiol. Educ. 29 (3) (2005) 151–159.
[39] M. Marullo, C. Zuccato, C. Mariotti, N. Lahiri, S.J. Tabrizi, S. Di Donato, E. Cattaneo, Expressed Alu repeats as a novel, reliable tool for normalization of real-time quantitative RT-PCR data, Genome Biol. 11 (1) (2010) R9.
[40] J. Vermeulen, K. De Preter, S. Lefever, J. Nuytens, F. De Vloed, S. Derveaux, J. Hellemans, F. Speleman, J. Vandesompele, Measurable impact of RNA quality on gene expression results from quantitative PCR, Nucleic Acids Res. 39 (9) (2011) e63.
[41] T. Nolan, R.E. Hands, S.A. Bustin, Quantification of mRNA using real-time RT-PCR, Nat. Protoc. 1 (3) (2006) 1559–1582.
[42] T. Nolan, R.E. Hands, W. Ogunkolade, S.A. Bustin, SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations, Anal. Biochem. 351 (2) (2006) 308–310.
[43] R.B. Ferns, J.A. Garson, Development and evaluation of a real-time RT-PCR assay for quantification of cell-free human immunodeficiency virus type 2 using a Brome Mosaic Virus internal control, J. Virol. Methods 135 (1) (2006) 162–168.
[44] J.F. Huggett, T. Novak, J.A. Garson, C. Green, S.D. Morris-Jones, R.F. Miller, A. Zumla, Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon, BMC Res. Notes 1 (2008) 70.
[45] R.G. Rutledge, Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications, Nucleic Acids Res. 32 (22) (2004) e178.
[46] R.G. Rutledge, D. Stewart, Critical evaluation of methods used to determine amplification efficiency refutes the exponential character of real-time PCR, BMC Mol. Biol. 9 (2008) 96.
[47] R.G. Rutledge, D. Stewart, A kinetic-based sigmoidal model for the polymerase chain reaction and its application to high-capacity absolute quantitative real-time PCR, BMC Biotechnol. 8 (2008) 47.
[48] R.G. Rutledge, D. Stewart, Assessing the performance capabilities of LRE-based assays for absolute quantitative real-time PCR, PLoS One 5 (3) (2010) e9731.
[49] A.N. Spiess, C. Feig, C. Ritz, Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry, BMC Bioinf. 9 (2008) 221.
[50] J.C.G.L.M. (JCGM), Evaluation of Measurement Data — Guide to the Expression of Uncertainty in Measurement (GUM), (2008).
[51] C. Kilkenny, W.J. Browne, I.C. Cuthill, M. Emerson, D.G. Altman, Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research, PLoS Biol. 8 (6) (2010) e1000412.
[52] A.M. Parfitt, M.K. Drezen, F.H. Glorieux, J.A. Kanis, H. Malluche, P.J. Meunier, S.M. Ott, R.R. Recker, Bone histomorphometry: standardization of nomenclature, symbols, and units. Report of the ASBMR Histomorphometry Nomenclature Committee, J. Bone Miner. Res. 2 (6) (1987) 595–610.