DOI: 10.1111/eci.12801

REVIEW

Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research

Stephen Bustin* and Tania Nolan†

*Postgraduate Medical Institute, Faculty of Medical Science, Anglia Ruskin University, Chelmsford, Essex, UK,
†Institute of Population Health, Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK

ABSTRACT

Poorly executed and inadequately reported molecular measurement methods are amongst the causes underlying the lack of reproducibility of much biomedical research. Although several high impact factor journals have acknowledged their past failure to scrutinise adequately the technical soundness of manuscripts, there is a perplexing reluctance to implement basic corrective measures. The reverse transcription real-time quantitative PCR (RT-qPCR) is probably the most straightforward measurement technique available for RNA quantification and is widely used in research, diagnostic, forensic and biotechnology applications. Despite the impact of the minimum information for the publication of quantitative PCR experiments (MIQE) guidelines, which aim to improve the robustness and the transparency of reporting of RT-qPCR data, we demonstrate that elementary protocol errors, inappropriate data analysis and inadequate reporting continue to be rife and conclude that the majority of published RT-qPCR data are likely to represent technical noise.

Keywords Gene expression, qPCR, quantification, reverse transcription.

Eur J Clin Invest 2017; 47 (10): 756–774

Background

Biomedical research is supported by immense sums of public and private funding, with the NIH alone investing over US$32 billion annually [1]. This generates tens of thousands of peer-reviewed research papers every year and drives the proliferation of new hypotheses, guides the direction of fresh research efforts, leads to the development of new treatments and so underpins further progress. However, the substantial increase in the number of scientific publications is also spurred by reasons other than the desire to communicate results to the scientific community, not least by the significance of researchers' publication output to their career progression [2], given the severe competition for tenured positions [3] and research funding [4]. The importance, if not the appropriateness [5], of using impact factors to rank the quality of research is well established and the most highly cited papers are published by a small number of prestigious journals [6]. Necessarily, there are hundreds of journals that share the remaining output and publish papers that can be equally significant, especially if they report technical innovations or report results that are, at the time, controversial. Inevitably then, regardless of impact factor, a key responsibility of the journal's editorial team is to ensure that there are procedures in place that serve as gatekeepers, ensuring that published results are broadly reproducible and, ideally, biologically relevant [7]. In theory, this is achieved through (i) editorial policies that ensure maximum transparency through the publication of accurate, comprehensive and explicit protocols and (ii) applying rigorous screening procedures, most often through the peer-reviewed process [8].

In practice, there are significant doubts about the validity of many research claims [9] in the context of a flawed research infrastructure that encourages disregard for responsible scientific process, regulation, transparency and reporting [10]. Confidence in quantitative measurements depends on a number of parameters, one of which is reproducibility [11]. Reproducibility incorporates both biological and technical variability, and as long ago as 1949, it was demonstrated that experimental test results can vary widely, even when performed by the same individual at the same time [12]. Since then, there have been numerous publications that highlight the problems of lack of reproducibility (reviewed in [13]) and the role journals play in failing to enforce their own editorial policies [14]. This, together with the fact that credibility and translation are only modestly correlated [15], explains why basic research findings are rarely adopted into clinical practice [16].

The problem of reproducibility

Data from a recent online survey run by *Nature* show that respondents believe that there is a significant (52%), or a slight (38%), crisis of reproducibility in the peer-reviewed, scientific literature, with 65% of respondents experiencing failure to even repeat their own results. Curiously, 73% still consider that at least half the papers in their field can be trusted [17]. This is borne out by the poor reproducibility of biomedical publications revealed in two studies [18,19], which has been associated with the decline in pharmaceutical research and development productivity [20]. Understanding the causes for this predicament is not straightforward, explanations are complex [21], and it has not been possible to identify one specific aspect of the published research that is responsible for this quandary [22]. In addition to the questionable relevance of many experimental models to the real world of human disease [23], the interpretation of experimental results is hampered by substantial biological [24] and technical [25] factors, as well as the deficiencies in the validation of experimental designs, [26] avoidable accumulation of practical errors [27] and inappropriate statistical analyses [28]. Together, these are the direct cause for the variability of published data that affect reproducibility, possibly going some way to explaining why, of the sciences, medicine and biology appear to have the lowest reproducibility [29]. Undoubtedly, lack of transparency in the Methods section of published papers represents a paramount, glaring and self-evident explanation, as full disclosure of protocols is essential to the integrity and utility of a research publication [30].

This subject has been discussed in numerous primary publications, reviews and editorials, with a range of causes evaluated in great detail using numerous examples and a range of solutions proffered. This includes prestigious and high impact factor journals, whose approach comprises assertions such as 'Sometimes replication is as important as discovery' [31] and calls for the publication of more detailed results from an experiment [32]. Indeed, Nature has acknowledged that editorial policies have contributed to 'failures in the reliability and reproducibility of published research' and admitted that journals 'compound' them by failing to 'exert sufficient scrutiny' and not publishing 'enough information for other researchers to assess results properly' [33], a message reaffirmed by many other Nature titles [34–40]. However, the suggested paths to action have remained largely untrodden, over the last 5 years peer-reviewed papers of dubious quality have continued to be published in low and high impact factor journals alike [14,41–45] and the scientific literature continues to be filled with thousands of papers that report results that are at best ambiguous and, at worst, simply wrong [13]. Regrettably, to date, the promises from many publishing houses to improve the transparency of reporting remain unfulfilled.

This is plainly demonstrated by considering the technical quality and completeness of reporting of real-time quantitative PCR (qPCR) and reverse transcription qPCR (RT-qPCR) assays, the two molecular techniques most widely used today. If the entire RT-qPCR workflow is carefully managed from the planning, optimisation and validation stages through the actual RT and qPCR cycling steps to data analysis and reporting steps, this apparently straightforward method can generate rapid, accurate, sensitive and cost-effective gene expression results [46]. However, the simplicity of the protocol and the ease with which results can be acquired obscure the pitfall of generating inappropriate or misleading results. Unfortunately, publications using both methods continue to be characterised by poor experimental design and lack of transparency of reporting of technical detail [47], factors recognised as contributors to poorly producible research [17], as well as by dubious analysis procedures [42]. This is despite the publication of the minimum information for the publication of quantitative PCR experiments (MIQE) guidelines, which define the minimum information necessary for evaluating the technical validity of published qPCR/RT-qPCR experiments [48], a further checklist for the optimisation and validation of qPCR assays [49] and additional guidelines for the validation of qualitative qPCR methods [50] as well as standards for the application of digital PCR [51]. In 2013, a review of publications from the Nature publishing group concluded that the reporting of essential information with regard to RT-qPCR protocols had actually deteriorated since the publication of the MIQE guidelines [13]. Whilst the guidelines strongly recommended inclusion of essential technical information, none of ten papers analysed reported this information. This means that the reader was provided with no information relating to RNA integrity, RNA purity, RT conditions or PCR efficiency and all of the publications used an inappropriate method of data analysis that was discredited fifteen years ago [52]. We have carried out a repeat analysis on twenty recent papers from Nature publishers, which paints essentially the same picture (Table 1). Ironically, transparency of reporting is negatively correlated with the impact factor of a journal [45]. Whilst these analyses relate to technical details and reporting inadequacies underlying RT-qPCR experiments, it is unlikely that this situation is unique to this single technique. Clearly, any change to this current state of affairs requires the editors of all journals to take seriously their obligation to ensure the technical validity of all papers they publish.

Techniques used for RNA expression studies

RNA profiling relies on a number of methods, some relatively simple to perform, others much more complex, some limited to

Table 1 Summary of technical information with regard to RT-qPCR provided by 20 publications

Publication	Year	RNA integrity	RT replicates	RT conditions	PCR conditions	PCR efficiency	Analysis	No of RG	RG validated	MIQE citation
154	2017	RIN ≥ 5	No information	No information	No information	No information	No information	1	No information	No
155	2015	No information	No information	Incomplete information	No information	No information	ΔΔCq	1	No information	No
156	2015	No information	No information	Incomplete information	No information	No information	ΔΔCq	1	No information	No
157	2016	No information	No information	No information	No information	No information	No information	1	No information	No
158	2016	No information	No information	No information	No information	No information	No information	No information	No information	No
159	2017	No information	No information	Incomplete information	No information	No information	ΔΔCq	1	No information	No
160	2017	No information	No information	Incomplete information	No information	No information	ΔΔCq	1	No information	No
161	2017	RT-qPCR methods directed to two citations, with no information or further direction provided								No
162	2017	No information	No information	Incomplete information	No information	No information	No information	1	No information	No
163	2016	RIN 9·5–10	No information	Incomplete information	No information	90–105%	ΔΔCq	1	Yes	Yes
164	2017	RIN ≥ 7	No information	Incomplete information	No information	No information	No information	1	Yes	No
165	2016	No information	No information	No information	No information	No information	ΔΔCq	1	No information	No
166	2016	No information	No information	Incomplete information	Yes	No information	No information	1	No information	No
167	2016	No information	No information	No information	No information	No information	ΔΔCq	Unclear 1 or 2	No information	No
168	2016	No information	No information	No information	No information	No information	ΔΔCq	1	No information	No
169	2016	No information	No information	No information	No information	No information	ΔΔCq	1	No information	No
170	2016	No information	No information	Yes	Yes	No information	No information	1	No information	No
171	2016	No information	No information	No information	No information	No information	ΔΔCq	1	No information	No

Table 1 *Continued*

Publication	Year	RNA integrity	RT replicates	RT conditions	PCR conditions	PCR efficiency	Analysis	No of RG	RG validated	MIQE citation
172	2016	No information	No information	No information	No information	No information	ΔΔCq	1	No information	No
173	2017	No information	No information	No information	No information	No information	ΔΔCq	1	No information	No

low throughput and others allowing the parallel processing of numerous samples.

1 PCR is an enabling technology for most of the methods currently in use and endpoint PCR analysed using gel electrophoresis continues to be used today, despite its limited quantification capacity. Awareness of variability problems associated with PCR has been long-standing, with the first report describing inconsistencies with replicate and serial specimens evaluated within and between laboratories as early as 1992 [53]. The lack of a theoretical understanding of the dynamic processes involved in PCR, especially with respect to the amplification of nonreproducible and/or unexpected amplification products, was also highlighted decades ago [54]. These observations and the resulting implications are largely disregarded.

2 qPCR and RT-qPCR were the first methods to promise rapid and easy quantification of nucleic acids and continue to be widely used, with a PubMed search for the terms 'reverse transcription real time PCR or RT-qPCR or qRT-PCR' returning nearly 5400 citations in 2016. The need for optimisation of the PCR step as an essential condition influencing the quantitative nature of a PCR was stressed from the beginning [55], yet qPCR is still referred to as a 'gold standard' [56,57] and continues to be widely perceived as easy to set up, undemanding to carry out, with results readily interpreted, even though adoption of that term was challenged many years ago [58]. In reality, there are numerous critical issues in the workflow that need to be addressed before biologically meaningful and trustworthy conclusions can be drawn [46] and there have been numerous publications questioning the method's true reproducibility, reliability and consistency (reviewed in [13]). However, 8 years after the publication of the MIQE guidelines, many papers reporting qPCR data continue to omit the most basic information, as shown in Table 1 and discussed below.

3 Microarrays allow much higher throughput than RT-qPCR, but due to the increase in complexity and time required to perform and analyse the data, this method is declining in popularity, with a PubMed search for the term 'microarray' returning around 1600 citations in 2016. As with RT-qPCR, there is evidence of initial lack of robustness of data, with the early focus on technical troubleshooting rather than the generation of reliable data of scientific significance [59] and clinical expression signatures dependent on whatever bioinformatics protocol was being applied [60]. Early cross-platform concordance has been reported as poor [61], although this may be a reflection of variable protocols and standards as more recent strict adherence to controlled standard operating procedures has resulted in higher intra- and interlaboratory reproducibility [62]. Again, it is not surprising that

there has been emphasis on the need for optimised protocols [63], proper quality control measures [64] together with creating consensus guidelines for use in clinical diagnostics [65]. Given the advent of whole transcriptome shotgun sequencing or RNA-seq, it is debatable how much longer microarrays will be in use, but they constitute a vast potential resource of information and have provided a legacy of published data, however reliable, or otherwise, they may be.

4 Whole transcriptome shotgun sequencing or RNA-seq is the technically most complex, expensive and time-consuming of all RNA profiling methods, making it less feasible to carry out replication or repeat experiments. Initially described as being a method where technical variability was too high to ignore [66], it is still unclear just how robust and reliable it is, although its use is increasing, reflected in the 2700 citations recorded by PubMed in 2016. However, the consensus is that RNA-seq offers the benefits of superior sensitivity [67] and high reproducibility [68], together with the power to identify novel transcriptomic features with potential use in clinical settings [69]. A multiplatform examination of RNA-seq data by the Sequencing Quality Control (SEQC) project found that relative, but not absolute, gene expression measurements can be measured accurately and reliably across laboratories and RNA-seq platforms [70] and that RNA-seq and microarray-based models were comparable in clinical endpoint prediction [71]. Current focus is on data analysis in general, where under controlled conditions reproducibility ranges from 60% to 93% [72]. Normalisation is a particular aspect where the application of different methodologies can generate discordant results [73–77]. There are also efforts underway to generate guidelines that are comparable to MIQE, for every stage of the RNA-seq workflow to safeguard the clinical relevance of the data and so ensuring reliable and robust clinical predictions [78]. However, it is worth bearing in mind that the reliability of much sequencing data depends on the robustness of the RT and PCR steps.

5 Two additional medium-throughput RNA profiling methods are the nCounter Analysis System from NanoString and the OpenArray System from Thermo Fisher. The former has the advantage of not requiring an RT step, the latter of reducing the amount of sample handling required. A comparison of mRNA abundance profiling by both technologies and RT-qPCR found that results were broadly comparable between the three methods. However, whereas NanoString and OpenArray results were very well correlated (R = 0·95), RT-qPCR data differed significantly from both (R = 0·48 with OpenArray and R = 0·55 with NanoString) [79]. The authors ascribe this to increased variability in the sample preparation caused by the larger amount of technician pipetting and plate set-up required for conventional RT-qPCR. Nevertheless, these results also demonstrate how results can be discordant even though techniques and analyses have been carried out in an appropriate manner. It also leaves open the possibility that it is the RT-qPCR results that are correct and that those obtained by the two other methods are not. This publication also serves as an excellent example of the amount of technical detail that should be included for the science to be evaluated.

This brief overview of the major RNA profiling methods currently in use illustrates the issues driving the vigorous debate about the reliability and biological/clinical relevance of many of the data published using RNA-based methods, and sheds light on the multifarious reasons underlying the realisation that these results are unreliable and in need of revision [9,10,16,26,80–87].

An accurate assessment of the extent and seriousness of irreproducibility has been hampered by the anecdotal nature of reports and a lack of quantitative data on specific failures to reproduce published biomedical research. This led to the establishment of the Reproducibility Project: Cancer Biology. The objective of this project was to replicate selected results from high-profile preclinical cancer biology research [88]. The initial findings from five studies have just been reported and indicate that two studies could replicate essential results reported by the original authors, one could not and two were uninterpretable because of technical issues. Interestingly, one of the replicated results [89] confirmed selective downregulation of MYC transcription by reverse transcription (RT) quantitative real-time polymerase chain reaction (qPCR) [90]. The original experiment compared MYC mRNA levels in control cells and cells treated for 1 or 8 h with a small molecule inhibitor and recorded a downregulation in treated cells to 6·9% and 8·8% of control cells. The replicate study recorded a downregulation to 14·6% and 12·6%, which is indeed similar but nevertheless amounts to a twofold difference. This is important because the changes in RNA levels reported by many publications using RT-qPCR data tend to be in that range, typically between 1·5- and fivefold.

A consequence of the superficially simple process of the RT-qPCR workflow is that on the one hand, many untrained and unskilled operators have adopted the technique. On the other hand, knowing where to look allows an experienced reader of a publication to identify problem areas and ascertain whether the reported results are likely to be real. Below follows a summary of the methodological issues that impede consistent interpretation of results, together with explicit examples that illustrate graphically why reported results are so frequently unreliable.

Pre-analysis

Relevant quality control measures are an essential preparation preceding any molecular biology-based experiment, including

ensuring reliable qPCR assays. There are numerous inhibitors that can be copurified during sample preparation and there are a variety of physical and chemical factors that degrade nucleic acids during the preparation process or subsequent storage. Hence, it is important that care is taken to determine whether there is a problem with purity and integrity and, for quantitative applications, this should be measurable and quantifiable [91]. In practice, however, most publications do not even mention sample purity, as measured by the absence of inhibition [45]. Yet the problems associated with inhibition have been well known for a long time [92] and their consequences are obvious. Even carrying out a general assay aimed at detecting inhibition [93] is probably not sufficient, as different PCRs are differentially susceptible to inhibition by substances copurified in nucleic acid extracts [94]. Although this has tremendous implications for microarray- and sequencing-based methods as well as molecular diagnostic applications, the phenomenon has been inadequately explored and remains ignored or at least unreported.

In principle, reliable quantification of RNA requires reasonably intact RNA [95]; the question is what constitutes 'reasonably' intact and how best to measure it. For microarray analysis, it has been suggested that comparisons are valid if they are done using samples of comparable RNA integrity [96] and a similar finding has been reported for RT-qPCR [97]. However, one needs to ask the question of how significant or correct results obtained from low-quality RNA are, especially if they are discordant from those obtained using high-quality RNA. Whilst it may be technically feasible to obtain data from such RNA, a demonstration of biological or clinical relevance is often missing. Since the original demonstration that assessing RNA quality by gel-based analysis of 18S and 28S rRNA is not sufficiently rigorous for quantitative applications [98], several methods have become available to assess RNA integrity. Some use a microcapillary electrophoretic separation system to provide a general assessment of overall RNA integrity through an RNA integrity number (RIN) [99], others measure mRNA degradation by comparing the amplification of different length amplicons of specific reference genes [100] or by evaluating the relative PCR amplification of 5′ and 3′ regions of cDNA obtained by oligo-dT priming [58]. Unfortunately, the majority of qPCR-based publications do not use or report data from such analyses and, at best, report spectrophotometric analysis data that are quite irrelevant for quantitative methods [101].

One suggestion is that total RNA with a RIN value of above five can be classified as good quality and above eight as perfect quality [95]. However, this is not an ideal categorisation as degradation is similar to inhibition in that it is not constant or a steady process but affects transcripts in different ways [102]. A serious consequence of this inconsistency is that comparisons of samples that differ in their purity or degree of degradation are likely to show differences where there are none or alternatively, mask any actual differences. This inherent instability compounds any inconsistencies arising from the intrinsic biological variability [58] and directly affects gene expression results and so distorts the clinical relevance of prognostic markers and risk classification [103]. Tellingly, even the adoption of rigorous quality criteria for RNA integrity assessment does not always prevent laboratory-specific effects leading to significant variability, although validated and standardised RT-qPCR assays can achieve accurate results [104].

It is worth noting that these considerations also apply to any qPCR experiments targeting DNA, where the choice of extraction method and the techniques used to quantify the extracted DNA, as well as the presence of inhibitors combine to result in measurement uncertainty [105].

The RT step

A scrutiny of the history of RT-PCR uncovers the unfortunate truth that the fundamental issues dogging the reliability of today's experiments were identified at an early stage, were constantly expounded and brought to the research community's attention, but continue to be disregarded by most. The potential of combining the RT and PCR steps to detect RNA targets was realised soon after the reporting of the first PCR experiment [106], with its use as a quantitative assay suggested soon afterwards [107]. However, it very quickly became clear that quantification of RNA was not quite as straightforward as applying the same procedures used for quantification of DNA as, for example, reliable reverse transcription of RNA for amplification by PCR depends on RNA concentration and priming methods [108] as well as the avoidance of extensive secondary structures at the primer binding sites [109,110]. Awkwardly, different RT enzymes are differentially affected by these secondary structures [111].

Misgivings with regard to the capacity of RT-PCR to distinguish small fold changes [112] were soon followed by reports of problems with lack of reproducibility of low copy number target detection [113], inhibition of the PCR by the RT enzyme [114], primer-dependent variability resulting in marked variation of RNA copy number [115] and a realisation that interlaboratory differences in the protocols used for sample preparation, RNA extraction and reverse transcription were responsible for discordance in published data involving endpoint RT-PCR [116]. The need for detailed quality assurance [117] and standardisation [118] at every step was highlighted well before the era of routine RT-qPCR usage, as were the challenges posed by the significant variability in RT and PCR efficiencies [119].

The coupling of RT with qPCR as a method for real-time data acquisition with RNA targets held out the promise of a

quantitative, as well as sensitive, accurate and high-throughput method [120] that might be more robust and reliable. However, substituting RT-PCR with RT-qPCR, in itself, does not control for RT variability [119]. Indeed, a careful reading of that first RT-qPCR paper reveals that the mean quantification cycles (Cq) of two RT-qPCR assays repeated on three separate days were reported as ranging from 29·3 to 29·6 and 17·4 to 18·3, respectively, *that is* the variability corresponded to 1·2- and 1·9-fold changes. Furthermore, there can be a priming-dependent lack of linearity of the RT step [121] and oligo-dT priming can lead to the generation of truncated cDNAs [122]. RNA yield at low template concentrations depends on the choice of RT enzyme [123] and accurate quantification of the amount of input RNA can be affected by variable amounts of contaminating DNA left after DNase treatment in the supposedly 'pure' RNA template [123]. What is more, the accuracy of RT-qPCR results is significantly affected by operator variability [123], which, by definition, introduces variation within and between laboratories.

The conclusive demonstration of the true variability of the RT step was provided by two key publications which ascertained that experimental variation in RT-qPCR is largely determined by the RT step. The first measured the properties of the RT reaction for five mRNA targets using different cDNA priming strategies [124]. Results not only confirmed the connection between priming strategy and both RT efficiency and RNA concentration, but also demonstrated that this dependency was target-specific, with the measured expression ratios of two genes varying by as much as 24-fold under different reaction conditions. Not surprisingly, the authors emphasised the need to optimise the RT step, carrying out at least duplicate RT reactions, adjusting the total RNA concentrations and always using the same reaction conditions. The second paper extended these findings to a comparison of the results obtained using eight different RT enzymes [125]. Disconcertingly, RT yields varied by up to 100-fold, dependent on which RT was used, and variation was also target-dependent. Unfortunately, the authors did not point out an obvious conclusion from both sets of results, namely that a quantitative technique generating inconsistent results with different reagents is neither effective nor useful. It is also worth reflecting on the fact that despite the central importance of these publications for understanding the properties of the RT reaction in mRNA quantification, their impact has been rather limited as citations are measured in the low hundreds compared to many thousands for key technical qPCR papers. More recently, these results were extended to demonstrate that the technical variability of the RT step is greater than the fold changes reported by many, if not most publications using RT-qPCR [126], again without much resonance in the research community.

Taken together, the conclusion inferred from these findings is that the RT step is intrinsically variable as it is carried out by different individuals that use variable amounts of RNA, a variety of cDNA priming methods and different RT enzymes together with a myriad of protocols. Awkwardly, the real-life implications of these findings are stark: a necessary consequence is that that RT-qPCR results reported by different laboratories are likely to be discordant and it is not an easy matter to ascertain which results are real and which ones are noise, especially if there is inadequate information on assay design parameters and experimental protocols. Clearly, this deduction is valid for any method that relies on the RT step, for example digital RT-PCR (RT-dPCR), where results are also both RT- and target-dependent [127], and RNA-seq, where technical variability has been found too high to ignore and has resulted in inconsistent detection of exons at low levels of coverage [66].

The PCR step

This step, usually described as a simple enzymatic reaction, is generally accepted to be the least troublesome and most reliable segment of a RT-qPCR assay, assuming that a few important quality criteria have been addressed. These include suitable assay design, empirical optimisation and validation of primers and amplicons and the absence of inhibitors in the template being amplified. Indeed, under those conditions PCR results are so reproducible that a traditional recommendation is to carry out replicates of the RT step and that it is acceptable to dispense with performing multiple qPCR replicates on each RT sample [128]. Alas, the robustness of the PCR step is not quite what it seems and there are several reasons for this.

1 Many published PCR assays are either inadequately designed and are unlikely to amplify only the intended target or the published information is incorrect, resulting in the publication of wrong primer or probe sequences. An example of the former is a qPCR assay for a truncated mRNA splice variant reported as a biomarker for predicting future metastasis in human metastatic colon, breast and hepatocellular cancers [129]. Analysis of the forward primer shows that it has sequence identity with 16 of 20 nucleotides present in the nontruncated mRNA, 15 of which are at the 3′ end, likely to result in effective nonspecific priming. An example of the latter is the reporting of a hydrolysis probe sequence that targets a region outside the PCR amplicon defined by the primers [130]. The surprise is that such basic mistakes can pass through the peer review system, as inadequate assays are easy to spot by a simple examination of the oligonucleotide sequences and assay location. These observations were only possible because the authors made their assay sequences available and therefore lend weight to the original MIQE [48] requirement for assay sequences to be published in full.

2 Optimal PCR performance is very much dependent upon appropriate template preparation methods [131] and there have now been more than 100 publications that have compared the performance of commercial extraction kits, and many more that of home-brew methods (recently reviewed in [105]. All of these comparisons show that it is important to choose extraction methods or commercial extraction kits with care, as they have a significant impact on the variability of subsequent PCR results. This strongly suggests that anyone contemplating carrying out qPCR assays should not just compare qPCR chemistries, reagents and instruments prior to settling on the optimal ones but also investigate nucleic extraction procedures, report the results and explain the final choice of extraction method.

3 The presumption that qPCR assays can reproducibly resolve small fold changes is misplaced. The specific manufacturer's claim of a twofold discriminatory power at 10 000 target copies is generalised to suggest that qPCR can always separate twofold differences. The concept is based upon standard instrument validation tests where β-actin copy numbers were measured in 36 replicates of 10 000 and 20 000 genome equivalents. As β-actin has multiple pseudogenes, the actual copy numbers measured may well have been greater than the genome equivalent. Therefore, it is inappropriate to extrapolate the instrument validation test to an assumption that all differences of twofold can be detected accurately, especially when running few, if any, replicates. As Fig. 1a and b shows, qPCR can indeed discriminate twofold differences and can do so reproducibly. However, at low target concentrations this is no longer possible, unless a large number of replicates are used and even then, there is a significant overlap of recorded Cqs. Furthermore, as Fig. 1c shows, it is clearly not possible to discriminate a 1·5-fold difference, something which is never claimed by manufacturers yet frequently reported. Just to quote two examples, Fig. 5 of a one paper reports 19 gene expression changes, of which 18 are < 1·5-fold [132] and Fig. 1 of another publication claims to be able to show 1·25-fold differences [133].

4 It is not much appreciated that all master mixes are not the same and that the use of modified polymerases, variable buffer additives and different Mg2+ concentrations can generate sufficiently variable results when different master mixes are used. Commercial master mixes vary in their ability to perform reliably across all instrument platforms [134], and indeed, some master mix/thermal cycler combinations are inadvisable [135]. Furthermore, the PCR step is known to introduce base composition bias in high-throughput sequencing libraries and DNA polymerase has been pinpointed as its principal cause [136], requiring modification and optimisation of standard amplification conditions. The extent and type of bias vary with different commercial master mixes, so much so that the avoidance of some master mixes may be required [137]. One reason for this bias is that sequences at the 3′ end of the primer as well as in the PCR amplicon itself appear to affect priming efficiencies of DNA polymerases, with different commercial enzymes having different preferences [138]. Hence, it is essential to understand that whilst replicates using the same master mix are very tight, repeat experiments using different master mixes can vary by at least and much as many of the reported fold differences. This is of particular relevance in the instances where Cq values only or Cq values relative to a single reference gene, a comparison referred to as ΔCq, are recorded and then published as fold differences. Figure 2 compares the expression of eight genes relative to a single reference gene after the amplification of the same cDNA with four master mixes. The variability recorded for each master mix, as evidenced by the error bars in Fig. 2b, is small, but considering that all reactions were carried out at the same time with the same instrument, the differences observed between master mixes are instructive. The variation in ΔCqs ranges from 3·4 to 5·1, which translates into fold differences from around 10- to about 30-fold (Fig. 2c). Without knowing the amplification efficiencies, it is impossible to resolve whether these differences could be caused by the three assays having different master mix-dependent priming or amplification efficiencies.

However, most RNA expression experiments compare the relative expression of markers between two samples, usually a test and a reference sample, calculating a fold difference, and the results shown in Fig. 3 show the technical noise expected from this type of comparison. The experiment involved using master mixes MM1 and MM2 to amplify cDNA prepared from two breast cancer RNA samples. The amplification plots for both master mixes are similar and the chosen reference gene, GAPDH, is expressed at similar levels in both RNA samples (Fig. 3a). However, results for genes of interest vary between two- and fivefold (Fig. 3b). Bearing in mind that these experiments were carried out by one researcher using RNA extracted by the same method, the same reagents and instruments, these fold differences constitute a minimum and could well increase if carried out by a different laboratory. Furthermore, the use of different primers may lead to smaller or larger differences.

5 RT-qPCR experiments can be carried out using separate RT reaction and PCR (two-step assays) or a combined reaction (one-step assay). The former tends to use oligo-dT and/or random primers, which provides more flexibility and enables the design of PCR-optimised target-specific primers. The latter minimises the hands-on time required to complete the assay and may be more sensitive depending on what assay is

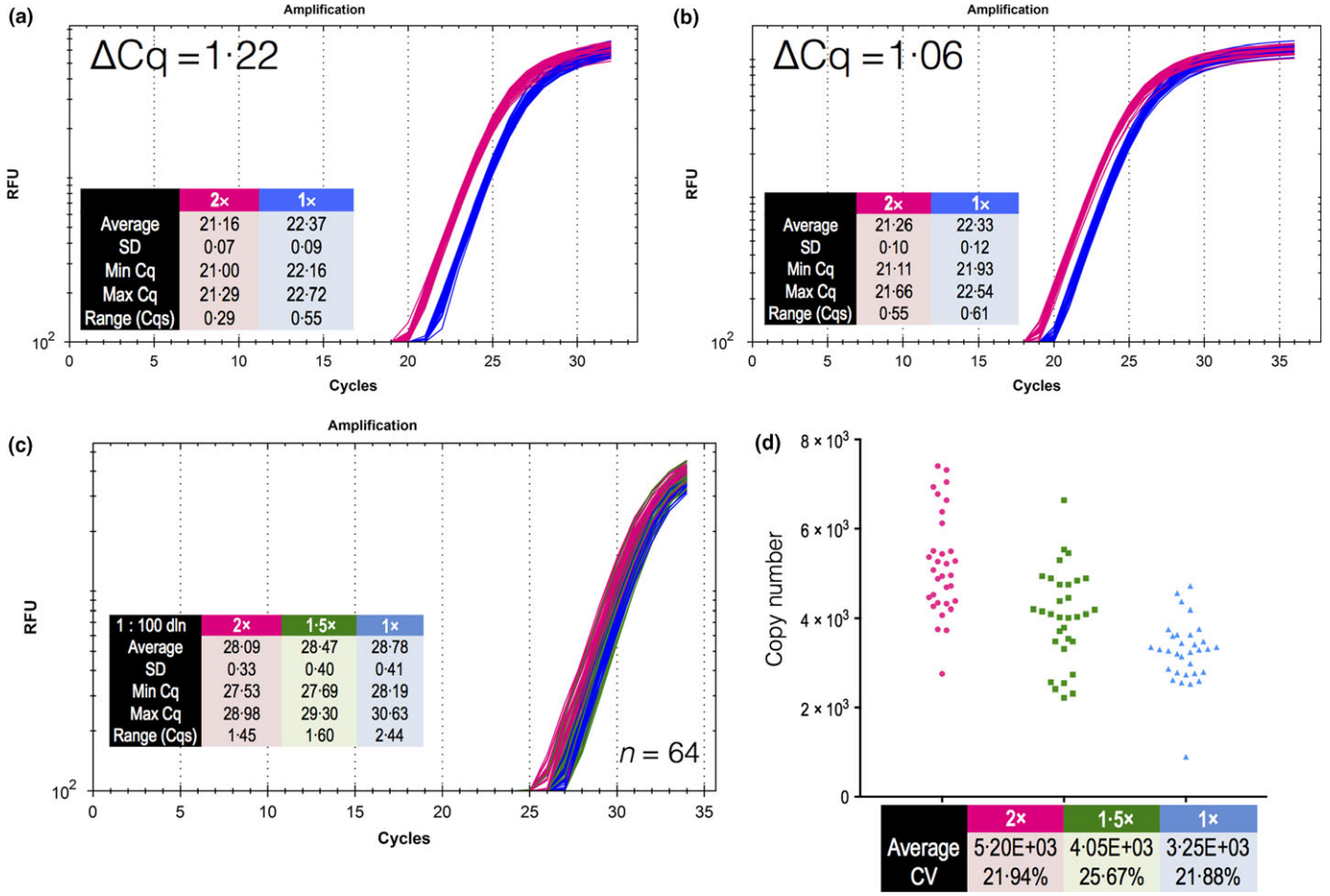

Figure 1 Resolution of a qPCR assay. (a) Amplification plots showing results from qPCRs carried out on genomic DNA, which was diluted twofold and subjected to a qPCR assay, with 48 replicates for each sample (neat sample is shown in red and diluted sample in blue). The observed ΔCq was 1·22, compared to a theoretical ΔCq of 1. (b) The assay was repeated on a different day with a fresh dilution of the sample and the ΔCq was 1·06. (c) The original sample was first diluted at a ratio of 1 in 100 and then again twofold. This time, additional qPCR assays were carried out using 1·5 times the amount of starting material (green amplification plots) from the twofold diluted sample, with 32 replicates for each sample. (d) Comparison of copy numbers calculated from a standard curve for the assay confirms the lack of resolution at this template concentration.

used. Comparisons between the two methods have concluded that results are broadly similar [139,140]. As always, it is dangerous to generalise, as this equivalence is not always observed. Figure 4 shows an example of the two approaches giving different results. Four mRNA targets were targeted in two RNA samples using either a two- or a one-step protocol. It is immediately apparent that the VDR results are significantly different for the two types of assay, with the one-step assay results suggesting either very low expression or primer dimer amplification. Four repeats of both assays gave the same result, indicating a problem with the robustness of the VDR data. An analysis of the secondary structure of the RNA (Fig. 4b) reveals extremely strong stem structures for all predictions throughout the template. One possible explanation therefore is that the target-specific primer used for the one-step assay could bind only inefficiently, whereas the use of random priming in the two-step assay allowed the RT to read through the sites subsequently used for PCR priming. However, both IGF1 and MMP3 primers also target secondary structure-rich RNA, so this cannot be the whole explanation. Furthermore, a comparison of fold changes reveals additional variability between the two methods (Fig. 4c). Whilst both generate concordant results for IGF1, they are discordant for MMP3. Without further detailed

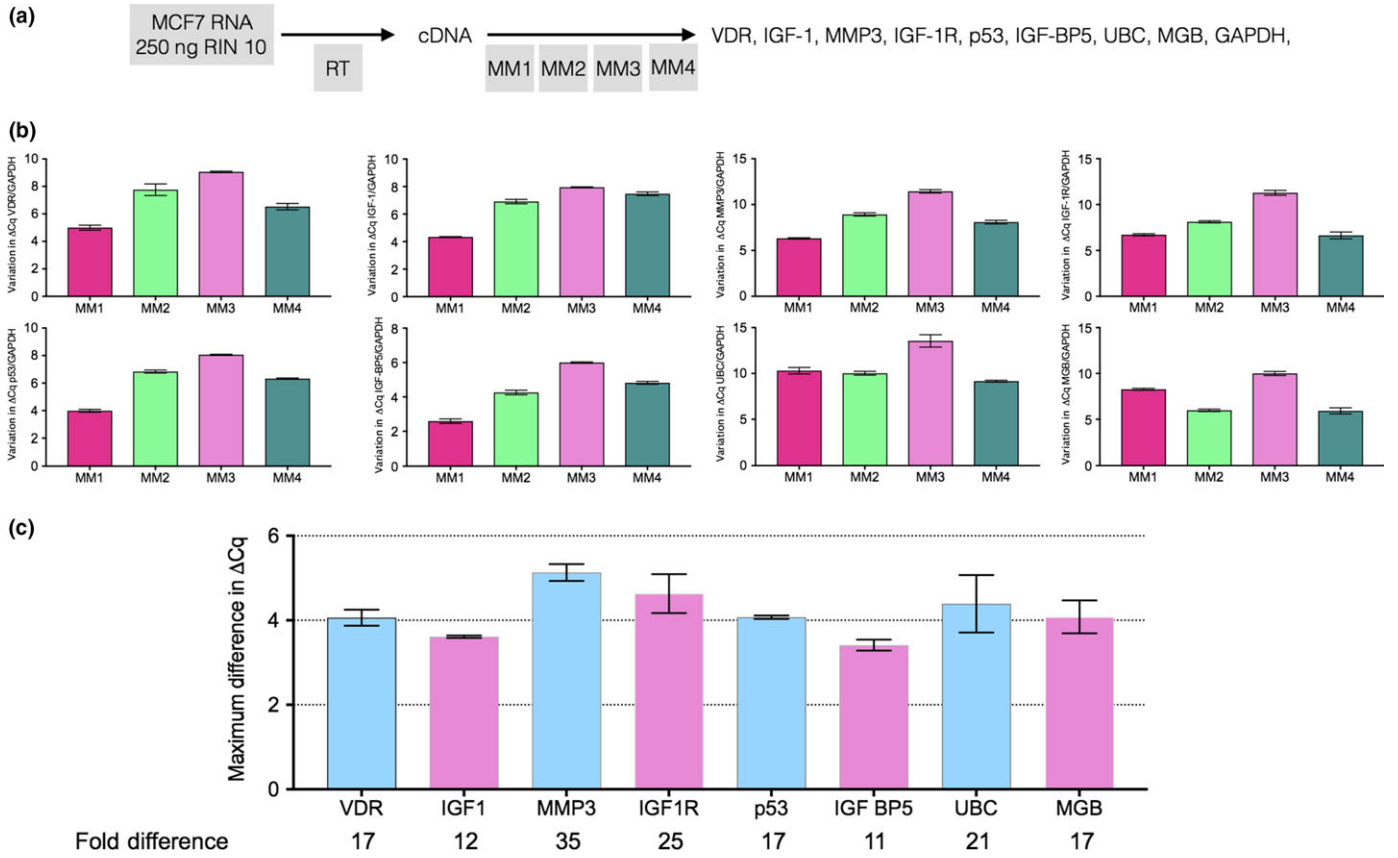

Figure 2 Master mix-dependent variability in cDNA quantification. (a) RNA extracted from a cell line with a RIN value of 10 was reverse-transcribed and nine targets were amplified with four master mixes (MM1-MM4) under identical, optimal conditions. (b) The expression of each target relative to that of GAPDH demonstrates the variation recorded in the ΔCq for each marker. (c) Plot of the maximum differences in ΔCq values and their translation into fold differences.

examination, it is impossible to ascertain why this is so. However, it reinforces the importance of careful primer design if the target is RNA.

6 One might assume that quality control of PCR reagents had been perfected after more than thirty years of the technique being at the forefront of molecular biology research. An early observation of maintaining PCR reagents at temperatures above the annealing temperature of the primers resulted in more specific amplification products [141] led the development of hot-start polymerases [142,143], which are now in common use for all varieties of PCR. Their use is especially important for high-throughput applications, where PCR reagents and templates are dispensed hours ahead of running the eventual PCR, giving any enzyme retaining polymerisation capacity plenty of time to generate spurious amplification products that could affect the sensitivity, specificity and reproducibility of the results. Unfortunately, many commercial hot-start polymerases appear not to have been tested by the suppliers for polymerase activity prior to thermal activation, as a recent report shows that 12 of 17 polymerases tested showed such activity when analysed by gel electrophoresis [144].

7 In the aforementioned survey [17], just under 60% of respondents posited that poor analysis was an important contributor to irreproducible research. Yet even for a method supposedly as simple as qPCR, there are numerous ways to analyse data. The need to address measurement uncertainty contributions made during the data analysis procedure was identified early on as important to minimise potential uncertainty in results due to subjective judgements [145]. The most commonly used method for comparing RNA transcript levels involves relative quantification against one or, preferably more, reference genes using the ΔΔCq method [146]. Although this approach has long been known to absolutely require knowledge of the PCR efficiencies of genes of interest and reference genes [147], repeated surveys show that most

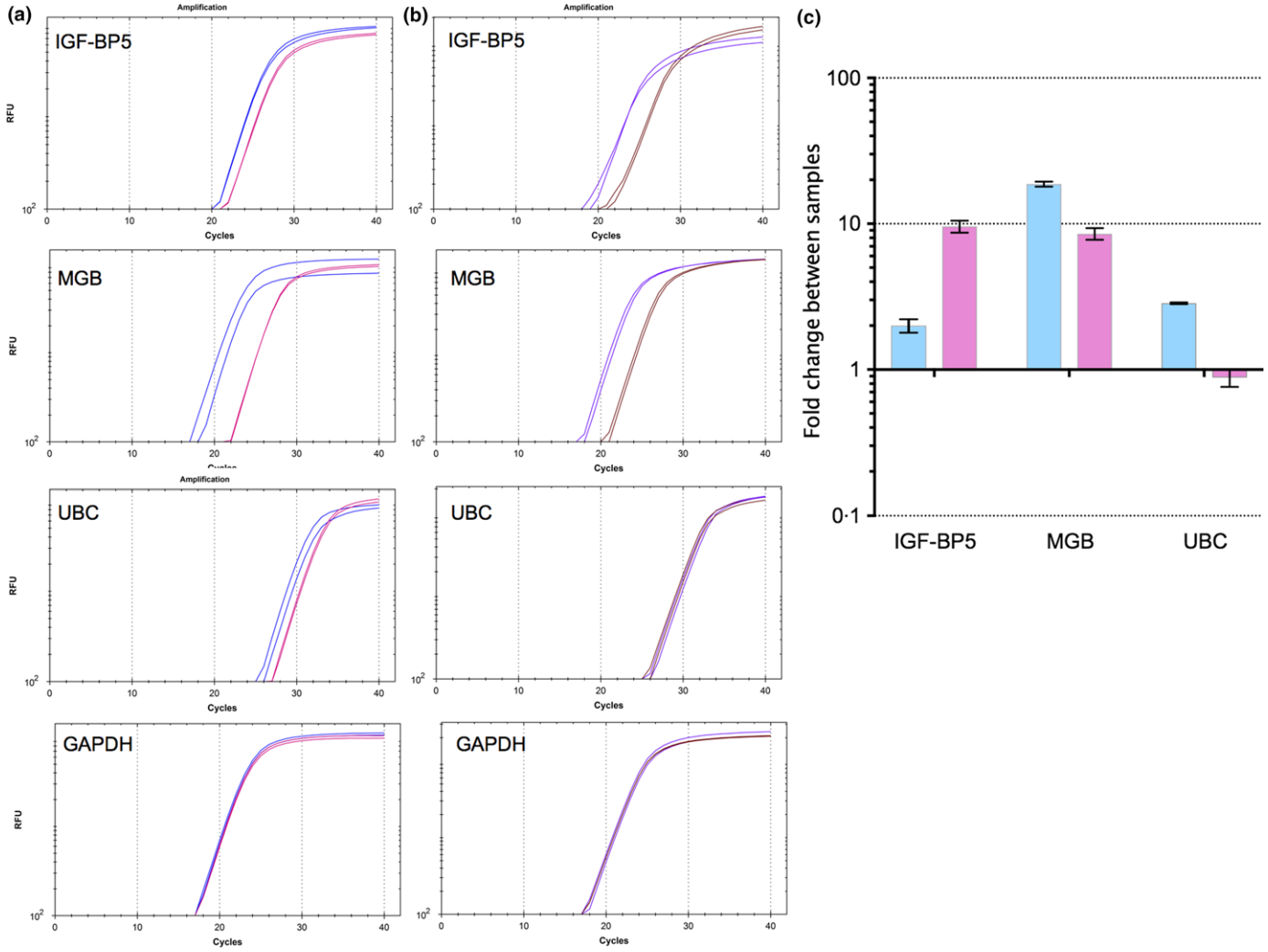

Figure 3 Technical noise of an RT-qPCR experiment. (a) Amplification plots of four targets amplified in duplicate from cDNAs prepared from two different high-quality RNAs (one shown in blue, the other in red) with RIN = 10 with MM1. (b) Amplification plots of the same targets amplified in duplicate from the same cDNAs (one shown in purple and the other in brown) with MM2. (c) The relative fold changes in expression of the three genes of interest (GOI) relative to that of GAPDH between the two samples is plotted (MM1 is blue and MM2 is pink).

researchers very rarely measure efficiency and assume it to be 100% [13,45,148], making any conclusions based on such experiments doubtful. This is still the case today, as the analysis in Table 1 reveals. About 8–10 years ago, a flurry of reports appeared, all attempting to introduce flexible, understandable and cost-effective data analysis and management tools [149]. These include qPCR-DAMS [150], Q-Anal [151], qpcR [152], QPCR [153] and qBase [154]. However, the various methods differ significantly in their performance, and whilst some generate reproducible and sensitive data, they do not always quantify cDNA with precision, especially if they bypass assumptions regarding PCR efficiency [155]. Indeed, the question of how to calculate PCR efficiency is another one of those many unsolved questions with respect to qPCR analysis, not least because recorded PCR efficiencies vary significantly across different instruments [156] and the performance of the different analysis tools cannot be easily compared as most use their own file format [157], raising the question of whether and how raw data should be submitted [158]. The standard curve approach is the most popular [159], but alternatives that estimate PCR efficiency based on the analysis of individual

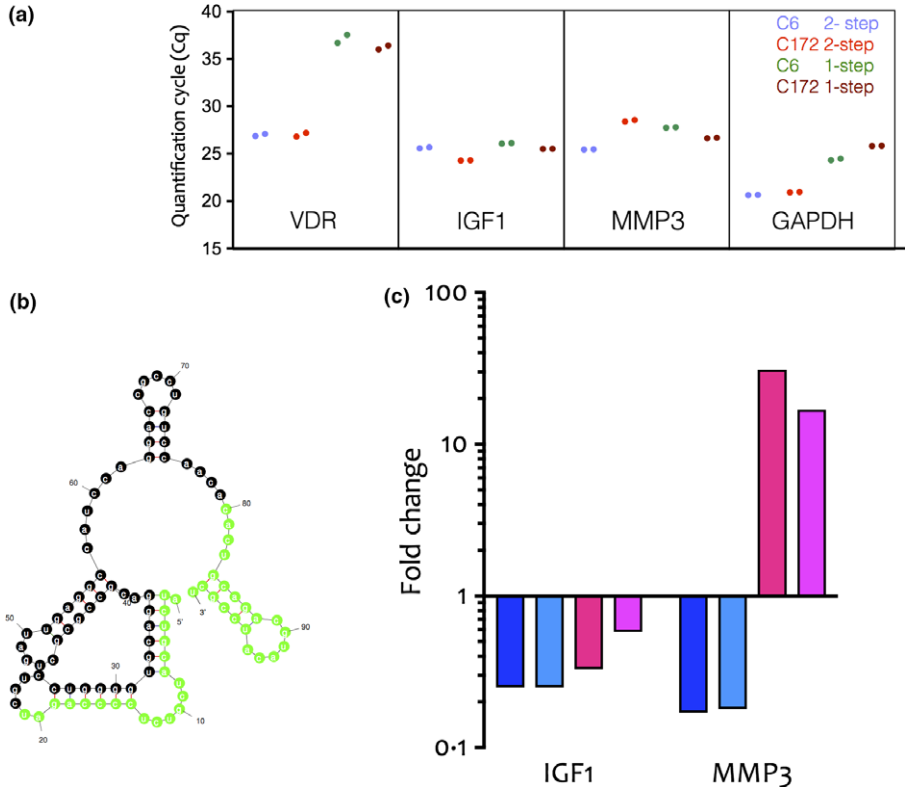

Figure 4 Comparison of one-step and two-step RT-qPCR. (a) Replicate Cq values recorded for one-step or two-step RT-qPCR assays carried out using the same RNA and PCR conditions and showing the lack of amplification for VDR with the one-step approach. (b) Secondary structure of VDR mRNA. (c) Comparison of the fold differences in expression of the two genes of interest (GOI) relative to that of GAPDH between two 1-step (Dark and light blue) and two 2-step (red and pink) protocols shows that results are concordant for one target but not the other.

amplification curves when using SYBR Green I detection chemistry claim to model PCR amplification more effectively [160–163].

These examples demonstrate the variability that can arise during the qPCR step and show that this is due to (i) inherent limitations of the ability of thermal cyclers to resolve small fold changes, especially at low template concentrations, (ii) variations between the master mixes caused by the use of different polymerases and additives to the buffers, (iv) differences between the performance of one-step and two-step RT-qPCR master mixes and (iii) insufficient attention by the manufacturers to quality control, resulting in partially active enzymes that could impinge on specificity and sensitivity. Notably, inappropriate RT-qPCR results are not simply 'false negative' or 'false positive', but are recorded as variable degrees or directions of relative changes in expression, expressed as fold differences. However, false positives are a problem if primers are not sufficiently specific, hence the need to validate them empirically and the strong recommendation to report their sequences [164,165].

Examples of poor practice

Two recent publications describe attempts to identify gene expression patterns associated with major depression disorder (MDD)-related behaviour or its treatment. The first used microarrays to identify differentially expressed transcripts in two animal models [166] and measured the relative expression of their human orthologs in the peripheral blood of patients diagnosed with MDD and control subjects using RT-qPCR. The authors concluded that they had established a clinically valid diagnostic panel of blood transcripts genes that can differentiate early-onset MDD from controls and MDD with or without comorbid anxiety. This study was extended in a follow-up study with a new cohort of patients [167], which examined the change in gene expression in patients with MDD prior to and in response to cognitive behaviour therapy (CBT) by RT-qPCR. As

the conclusions of the two studies are based on interpretation of RT-qPCR data, it is germane to scrutinise both publications for technical validity and analytical soundness.

As is usual, there is little transparency of reporting of technical information: (i) the first paper [166] does not mention either RNA integrity or purity assessment, the second refers to spectrometric analysis but provides no information on sample quality, (ii) there is inadequate reporting of RT conditions, (iii) a single reference gene is used with no evidence that this is expressed stably across all conditions, (iv) there are mistakes in the published primer sequences, (v) primers do not detect all splice variants, and (vi) one assay is nonexon spanning, but no controls are provided to show the absence of contaminating genomic DNA.

However, the most troubling issue is concerned with the analysis and interpretation of the data. Fortunately, the authors have included some limited analysis data in each publication. These have been collated in Table 2. In detail, we note the following errors:

1 The authors record ΔCq values between the genes of interest and the reference gene, but there is no evidence that they evaluated PCR efficiency. As discussed previously, this is enormously concerning, especially in the light of the minute differences in expression being reported.

2 A comparison of ΔCqs from the two studies reveals a striking difference between the 2012 and 2014 studies, with all transcripts recording higher values in the later paper compared with the earlier one (Table 2). The average Cq difference for all markers in the MDD groups was 4·81 (or 28-fold) and 5·5 (or 45-fold) in the control groups. However, this masks large differences between the individual markers, with FAM46A recording the same results and CADM1/IGSF4A differing by sevenfold. As the methods sections of both studies are virtually the same, these differences strongly suggest a technical difference between the studies. However, as there is inadequate information with regard to the protocols used, it is impossible to discern whether there were potentially significant differences in RNA quality, conversion to cDNA, qPCR efficiency or analysis procedures or a combination of these parameters.

3 The ΔΔCq method of calculation has been applied incorrectly, as the authors report negative fold changes and it is

Table 2 Genes and ΔCq values extracted from two publications analysing MDD [166,167]

Gene	2012 MDD	2014 MDD	Cq difference	2012 ND	2014 ND	Cq difference
ATP11C	15·26	19·22	3·96	14·92	19·20	4·28
CD59	14·65	19·98	5·33	13·98	19·89	5·91
FAM46A	15·31	19·65	4·34	15·62	19·96	4·34
CADM1/IGSF4A	18·42	22·22	3·80	16·50	23·05	6·55
MAF	18·12	22·00	3·88	17·70	22·08	4·38
MARKS	13·52	16·85	3·33	13·29	17·41	4·12
NAGA	13·69	17·63	3·94	13·90	17·85	3·95
RAPH1	17·11	21·73	4·62	15·70	22·80	7·10
TLR7	15·04	19·53	4·49	15·43	19·89	4·46
ZNF291	16·17	23·55	7·38	16·16	23·20	7·04
AMFR	13·98	19·28	5·30	13·57	19·27	5·70
CDR2	15·55	20·60	5·05	14·88	20·49	5·61
CMAS	17·32	22·41	5·09	15·53	22·98	7·45
DGKA	10·31	16·69	6·38	11·24	17·41	6·17
PSME1	11·66	17·17	5·51	11·20	17·55	6·35
PTP4A3	15·80	20·79	4·99	14·95	21·03	6·08
SLC4A1	12·69	17·08	4·39	13·19	17·22	4·03
Average			4·81			5·50
SD			1·01			1·20

The two studies use different cohorts of patients and controls but have similar clinical characteristics. MDD, major depression disorder.

impossible to produce a negative value from the Livak algorithm ($2^{-\Delta\Delta Cq}$). The actual ΔΔCq values between MDD and control groups for each marker are very small in both studies, suggesting minimal differences in expression between the two groups. For example, the ΔΔCq values for ATP11C were 0·79 and 0·99 in the earlier and later reports, respectively. The correct changes have been recalculated for each data set and are summarised in Fig. 5. None of the data show any genes with significant fold changes, beyond that which could be explained by experimental variation. In fact, RAPH1, the gene highlighted as potentially diagnostic and for which a patent has been applied for shows a change in direction of expression between the two studies.

Guidance

There are a number of obvious steps that researchers carrying out RT-qPCR experiments and reviewers or readers of a publication can take to determine how reliable any results and subsequently reported data are likely to be. The aim is to ensure the experimental design and workflow controls as much as possible for biological variability, caused by inherent differences between individual organisms, tissues or cells and technical variability originating from experimental elements. The former is best addressed through the use of three or so biological replicates per experimental sample whenever feasible, the latter using three technical replicates per reaction.

1 *In silico* verification of primer and probe sequences.

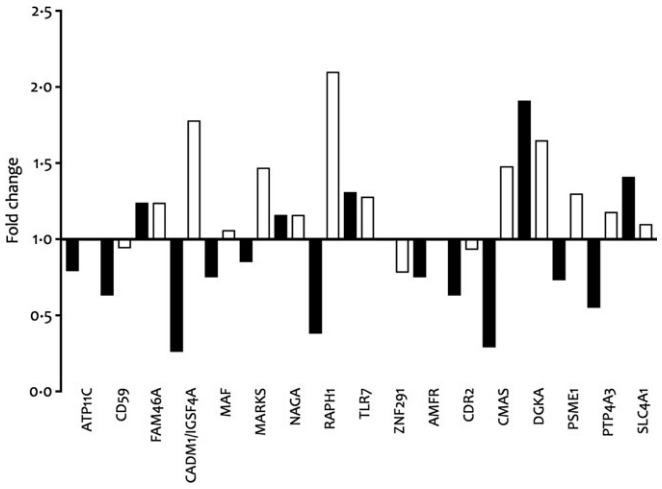

Figure 5 Recalculated fold changes for the targets claimed to show significant differences between patients with MDD and normal controls. The black bars show the values recorded by the earlier publication [166].

1.1 As the specificity and efficiency of any PCR assay are critically dependent on the primers and probes, if used, these need to be carefully checked. Unfortunately, there are frequent mistakes, including simple typographical errors, confusion about the oligonucleotide sequence orientation, primers that lack specificity, bind to multiple targets, have 3′ ends that target SNP sites, are not at all or insufficiently intron spanning or probes that bind to targets outside the PCR amplicon. For one tube combined RT-qPCR assays, it is also important to establish that the cDNA primers do not target strong secondary structures of the RNA, as this results in variable, low efficient priming by the reverse transcriptase.

2 Establish the purity (absence of inhibitors) and determine the quality (integrity) of nucleic acid samples following their extraction.

2.1 The absence of inhibition can be tested either by diluting a sample and comparing the Cqs or by using a reporter such as the SPUD assay. The former method requires more sample and reagents, but allows an assessment of any potential inhibition on the targets of interest. The latter is less labour-intensive but may miss any target-specific inhibition. Sample quality is best measured using a microcapillary system or a 3′-5′ assay. Reporting an $OD_{260:280}$ ratio is meaningless for experiments aiming to quantify accurately small differences in RNA concentrations. Inadequate reporting means that it is never clear whether authors actually check for the presence of inhibitors and ascertain an RNA sample's integrity or just do not report their results.

3 The RT step should be carried out in triplicate if at all possible, whereas it is sufficient to run a single qPCR for each RT reaction.

3.1 Most reported fold changes are relatively small and the variability of the RT step, especially when using low concentrations of RNA, has been demonstrated to be greater than many of the reported fold changes. Yet there are virtually no publications that follow this recommendation. The qPCR step, on the other hand, is highly reproducible, if the same master mix is used by a skilful operator and so multiple reactions are less informative. Alas, a change of master mix can generate different results, especially if the expression changes are marginal.

4 The efficiency of each individual PCR assay must be ascertained.

4.1 The validity of a quantitative result cannot be determined reliably and reproducibly without knowing whether the PCR assays were efficient and incorporating any variabilities into fold change calculations. PCR efficiency can be calculated from standard curves or using published, often free-to-use, software. Most qPCR instruments provide an option to calculate PCR efficiencies automatically. The use of the ΔΔCq method in the absence of information on PCR efficiency is always a clear indicator that the data are suspect. If researchers were to address only this one element of the RT-qPCR workflow, this would arguably result in the biggest improvement to the accuracy of gene expression analyses.

5 The use of multiple, stable and validated reference genes is essential for RT-qPCR data normalisation as it removes variation due to inherent technical or experimentally induced variation.

5.1 Normalisation against a single, unvalidated reference gene to report small fold changes has long been known to invalidate quantitative results and results in erroneous normalisation up to threefold and 6·4-fold in 25% and 10% of the cases, respectively, with sporadic cases showing error values above 20% [52]. In contrast, most reports deal with fold changes that range from 1·5-fold to about eightfold, all of which are definitely included in this category. Importantly, depending on the amount of target being quantified, fold changes less than three- to fivefold are likely to be unreliable.

6 Cq data should be submitted as supplementary information, either in a spreadsheet or directly exported in RDML format.

Conclusions

The ostensible simplicity of RT-qPCR conceals a complexity that extends to every step of the workflow and there is a long history of publications analysing, untangling and clarifying the issues that must be confronted to ensure the appropriateness of assay design, robustness of the protocol and finally the suitability of analysis methods. On the other hand, there are numerous publications that have been highlighting serious concerns with the relevance, reproducibility and practicability of biomedical research. Considering the key enabling roles of the RT and PCR steps in other molecular technologies, it is reasonable to put the two issues together and conclude that there is a link between the two. It is shocking that despite the easy availability of this information, the quality of execution of a technique that has been in use for 24 years remains at such a low standard. Either its users are ignorant of the complications involved in ensuring reproducibility or they deliberately ignore the wealth of advice and knowingly publish inadequate results. It is also clear that the availability of guidelines has done very little to improve the quality of published data based on the use of RT-qPCR and begs the question that if a technique where inadequacies are easily detected is as badly executed as this, what must the situation be like for the much more complex technologies based on it. The effects of this negligence by the research community to heed the many warnings that have been issued over the years is enhanced by the failure shown by the editors of most journals to implement their own instructions for authors, despite publishing editorials and seemingly supporting initiatives aimed at improving this situation. This unwillingness to take remedial action in general and for RT-qPCR in particular means that much taxpayers' money and donors' generosity continue to be wasted.

Address
Postgraduate Medical Institute, Faculty of Medical Science, Anglia Ruskin University, Bishop Hall Lane, Chelmsford, Essex, CM1 1SQ, UK (S. Bustin); Institute of Population Health, Faculty of Medical and Human Sciences, University of Manchester, M13 9PL, UK (T. Nolan).

Correspondence to: Stephen Bustin, Postgraduate Medical Institute, Faculty of Medical Science, Anglia Ruskin University, Bishop Hall Lane, Chelmsford, Essex, CM1 1SQ, UK. Tel.: 01245 68 4804; e-mail: Stephen.bustin@anglia.ac.uk

Received 14 March 2017; accepted 7 August 2017

References

1 Health NIO. Budget - Research for the people.
2 Dinis-Oliveira RJ, Magalhães T. The inherent drawbacks of the pressure to publish in health sciences: good or bad science. *F1000Res* 2015;**4**:419.
3 Stephan P. Research efficiency: Perverse incentives. *Nature* 2012;**484**:29–31.
4 Alberts B, Kirschner MW, Tilghman S, Varmus H. Rescuing US biomedical research from its systemic flaws. *Proc Natl Acad Sci USA* 2014;**111**:5773–7.
5 Coleman R. Impact factors: use and abuse in biomedical research. *Anat Rec* 1999;**257**:54–7.
6 Ioannidis JP. Concentration of the most-cited papers in the scientific literature: analysis of journal ecosystems. *PLoS ONE* 2006;**1**:e5.
7 Bustin S, Huggett J. Reproducibility of biomedical research – The importance of editorial vigilance. *Biomol Det Quant* 2017;**11**:1–3.
8 Laine C, Goodman SN, Griswold ME, Sox HC. Reproducible research: moving toward research the public can really trust. *Ann Intern Med* 2007;**146**:450–3.

9 Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;**2**:e124.
10 Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP *et al*. Biomedical research: increasing value, reducing waste. *Lancet* 2014;**383**:101–4.
11 Plant AL, Locascio LE, May WE, Gallagher PD. Improved reproducibility by assuring confidence in measurements in biomedical research. *Nat Methods* 2014;**11**:895–8.
12 Narins SR. Studies on scratch tests in atopic dermatitis, and comparison with the intracutaneous method; the lack of reproducibility of results with the scratch tests. *J Invest Dermatol* 1949;**12**:49–60.
13 Bustin SA. The reproducibility of biomedical research: Sleepers awake. *Biomol Detect Quantif* 2014;**2**:35–42.
14 Bustin SA, Nolan T. Improving the reliability of peer-reviewed publications: We are all in it together. *Biomol Detect Quantif* 2016;**7**: A1–5.
15 Ioannidis JP. Evolution and translation of research findings: from bench to where? *PLoS Clin Trials* 2006;**1**:e36.
16 Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. *Am J Med* 2003;**114**:477–84.
17 Baker M. Is gthere a reproducibility crisis? *Nature* 2016;**533**:452–4.
18 Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 2011;**10**:712.
19 Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature* 2012;**483**:531–3.
20 Kannt A, Wieland T. Managing risks in drug discovery: reproducibility of published findings. *Naunyn Schmiedebergs Arch Pharmacol* 2016;**389**:353–60.
21 Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res* 2015;**116**:116–26.
22 Jarvis MF, Williams M. Irreproducibility in preclinical biomedical research: perceptions, uncertainties, and knowledge gaps. *Trends Pharmacol Sci* 2016;**37**:290–302.
23 Horrobin DF. Modern biomedical research: an internally self-consistent universe with little contact with medical reality? *Nat Rev Drug Discov* 2003;**2**:151–4.
24 Svendsen O, Hansen Kornerup A. Biological variation, reproducibility, and predictability of experimental research in animals. In: Thomsen H, Muller RN, Mattrey RF, editors. *Trends in Contrast Media*. Berlin Heidelberg: Springer; 1999: pp 31–41.
25 Bustin SA, Dorudi S. Molecular assessment of tumour stage and disease recurrence using PCR-based assays. *Mol Med Today* 1998;**4**:389–96.
26 Ioannidis JP, Khoury MJ. Improving validation practices in "omics" research. *Science* 2011;**334**:1230–2.
27 Casadevall A, Steen RG, Fang FC. Sources of error in the retracted scientific literature. *FASEB J* 2014;**28**:3847–55.
28 Thiese MS, Arnold ZC, Walker SD. The misuse and abuse of statistics in biomedical research. *Biochem Med (Zagreb)* 2015;**25**:5–11.
29 Baker M, Dolgin E. Reproducibility project yields muddy results. *Nature* 2017;**541**:269–70.
30 Anonymous. Reproducibility in research. [editorial]. *Dis Model Mech* 2011;**4**:279.
31 Anonymous. Further confirmation needed. [editorial]. *Nat Biotechnol* 2012;**30**:806.
32 Anonymous. Must try harder. [editorial]. *Nature* 2012;**483**:509.
33 Anonymous. Announcement: Reducing our irreproducibility. *Nature* 2013;**496**:398.
34 Anonymous. Raising reporting standards. *Nat Cell Biol* 2013;**15**:443.
35 Anonymous. Raising standards. *Nat Biotechnol* 2013;**31**:366.
36 Anonymous. Raising standards. *Nat Med* 2013;**19**:508.
37 Anonymous. Raising standards. *Nat Struct Mol Biol* 2013;**20**:533.
38 Anonymous. Raising standards. *Nat Genet* 2013;**45**:467.
39 Anonymous. Raising standards. *Nat Neurosci* 2013;**16**:517.
40 Anonymous. Raising standards. *Nat Immunol* 2013;**14**:415.
41 Abdel Nour AM, Azhar E, Damanhouri G, Bustin SA. Five years MIQE guidelines: the case of the Arabian countries. *PLoS ONE* 2014;**9**:e88266.
42 Dijkstra JR, van Kempen LC, Nagtegaal ID, Bustin SA. Critical appraisal of quantitative PCR results in colorectal cancer research: Can we rely on published qPCR results? *Mol Oncol* 2014;**8**:813–8.
43 Bustin S. Transparency of reporting in molecular diagnostics. *Int J Mol Sci* 2013;**14**:15878–84.
44 Bustin SA, Murphy J. RNA biomarkers in colorectal cancer. *Methods* 2013;**59**:116–25.
45 Bustin SA, Benes V, Garson J, Hellemans J, Huggett J, Kubista M *et al*. The need for transparency and good practices in the qPCR literature. *Nat Methods* 2013;**10**:1063–7.
46 Derveaux S, Vandesompele J, Hellemans J. How to do successful gene expression analysis using real-time PCR. *Methods* 2010;**50**:227–30.
47 Garson JA, Huggett JF, Bustin SA, Pfaffl MW, Benes V, Vandesompele J *et al*. Unreliable real-time PCR analysis of human endogenous retrovirus-W (HERV-W) RNA expression and DNA copy number in multiple sclerosis. *AIDS Res Hum Retroviruses* 2009;**25**:377–8. Author reply 379.
48 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.
49 Raymaekers M, Smets R, Maes B, Cartuyvels R. Checklist for optimization and validation of real-time PCR assays. *J Clin Lab Anal* 2009;**23**:145–51.
50 Broeders S, Huber I, Grohmann L, Berben G, Taverniers I, Mazzara M *et al*. Guidelines for validation of qualitative real-time PCR methods. *Trends Food Sci Technol* 2014;**37**:115–26.
51 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.
52 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A *et al*. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;**3**:0034.1–0034.11.
53 Busch MP, Henrard DR, Hewlett IK, Mehaffey WF, Epstein JS, Allain JP *et al*. Poor sensitivity, specificity, and reproducibility of detection of HIV-1 DNA in serum by polymerase chain reaction. The Transfusion Safety Study Group. *J Acquir Immune Defic Syndr* 1992;**5**:872–7.
54 Schierwater B, Metzler D, Krüger K, Streit B. The effects of nested primer binding sites on the reproducibility of PCR: mathematical modeling and computer simulation studies. *J Comput Biol* 2009;**3**:235–51.
55 Ferre F. Quantitative or semi-quantitative PCR: reality versus myth. *PCR Methods Appl* 1992;**2**:1–9.
56 Yoo JE, Lee C, Park S, Ko G. Evaluation of various real-time reverse transcription quantitative PCR assays for norovirus detection. *J Microbiol Biotechnol* 2017;**27**:816–24.

57 Alikian M, Whale AS, Akiki S, Piechocki K, Torrado C, Myint T *et al.* RT-qPCR and RT-digital PCR: a comparison of different platforms for the evaluation of residual disease in chronic myeloid leukemia. *Clin Chem* 2017;**63**:525–31.
58 Nolan T, Hands RE, Bustin SA. Quantification of mRNA using real-time RT-PCR. *Nat Protoc* 2006;**1**:1559–82.
59 Shi L, Tong W, Goodsaid F, Frueh FW, Fang H, Han T *et al.* QA/QC: challenges and pitfalls facing the microarray community and regulatory agencies. *Expert Rev Mol Diagn* 2004;**4**:761–77.
60 Miklos GL, Maleszka R. Microarray reality checks in the context of a complex disease. *Nat Biotechnol* 2004;**22**:615–21.
61 Sato F, Tsuchiya S, Terasawa K, Tsujimoto G. Intra-platform repeatability and inter-platform comparability of microRNA microarray technology. *PLoS ONE* 2009;**4**:e5540.
62 Ach RA, Floore A, Curry B, Lazar V, Glas AM, Pover R *et al.* Robust interlaboratory reproducibility of a gene expression signature measurement consistent with the needs of a new generation of diagnostic tools. *BMC Genom* 2007;**8**:148.
63 Li L, Roden J, Shapiro BE, Wold BJ, Bhatia S, Forman SJ *et al.* Reproducibility, fidelity, and discriminant validity of mRNA amplification for microarray analysis from primary hematopoietic cells. *J Mol Diagn* 2005;**7**:48–56.
64 Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC *et al.* The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006;**24**:1151–61.
65 Staal FJ, Cario G, Cazzaniga G, Haferlach T, Heuser M, Hofmann WK *et al.* Consensus guidelines for microarray gene expression analyses in leukemia from three European leukemia networks. *Leukemia* 2006;**20**:1385–92.
66 McIntyre LM, Lopiano KK, Morse AM, Amin V, Oberg AL, Young LJ *et al.* RNA-seq: technical variability and sampling. *BMC Genom* 2011;**12**:293.
67 Xu J, Thakkar S, Gong B, Tong W. The FDA's experience with emerging genomics technologies-past, present, and future. *AAPS J* 2016;**18**:814–8.
68 Costa C, Giménez-Capitán A, Karachaliou N, Rosell R. Comprehensive molecular screening: from the RT-PCR to the RNA-seq. *Transl Lung Cancer Res* 2013;**2**:87–91.
69 Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat Rev Genet* 2016;**17**:257–71.
70 SEQC/MAQC-III C. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat Biotechnol* 2014;**32**:903–14.
71 Xu J, Gong B, Wu L, Thakkar S, Hong H, Tong W. Comprehensive Assessments of RNA-seq by the SEQC Consortium: FDA-Led Efforts Advance Precision Medicine. *Pharmaceutics* 2016;**8**:8. https://doi.org/10.3390/pharmaceutics8010008.
72 Łabaj PP, Kreil DP. Sensitivity, specificity, and reproducibility of RNA-Seq differential expression calls. *Biol Direct* 2016;**11**:66.
73 Gupta R, Dewan I, Bharti R, Bhattacharya A. Differential Expression Analysis for RNA-Seq Data. *ISRN Bioinform* 2012;**2012**:817508.
74 Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* 2012;**13**:204–16.
75 Rau A, Marot G, Jaffrézic F. Differential meta-analysis of RNA-seq data from multiple studies. *BMC Bioinformatics* 2014;**15**:91.
76 Shin H, Shannon CP, Fishbane N, Ruan J, Zhou M, Balshaw R *et al.* Variation in RNA-Seq transcriptome profiles of peripheral whole blood from healthy individuals with and without globin depletion. *PLoS ONE* 2014;**9**:e91041.
77 Li P, Piao Y, Shon HS, Ryu KH. Comparing the normalization methods for the differential analysis of Illumina high-throughput RNA-Seq data. *BMC Bioinformatics* 2015;**16**:347.
78 Buschmann D, Haberberger A, Kirchner B, Spornraft M, Riedmaier I, Schelling G *et al.* Toward reliable biomarker signatures in the age of liquid biopsies - how to standardize the small RNA-Seq workflow. *Nucleic Acids Res* 2016;**44**:5995–6018.
79 Prokopec SD, Watson JD, Waggott DM, Smith AB, Wu AH, Okey AB *et al.* Systematic evaluation of medium-throughput mRNA abundance platforms. *RNA* 2013;**19**:51–62.
80 Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM *et al.* How to increase value and reduce waste when research priorities are set. *Lancet* 2014;**383**:156–65.
81 Djulbegovic B, Hozo I. When should potentially false research findings be considered acceptable? *PLoS Med* 2007;**4**:e26.
82 Ioannidis JP. Is molecular profiling ready for use in clinical decision making? *Oncologist* 2007;**12**:301–11.
83 Ioannidis JP. Limitations are not properly acknowledged in the scientific literature. *J Clin Epidemiol* 2007;**60**:324–9.
84 Ioannidis JP. Molecular evidence-based medicine: evolution and integration of information in the genomic era. *Eur J Clin Invest* 2007;**37**:340–9.
85 Ioannidis JP. Expectations, validity, and reality in omics. *J Clin Epidemiol* 2010;**63**:945–9.
86 Ioannidis JP. How to make more published research true. *PLoS Med* 2014;**11**:e1001747.
87 Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D *et al.* Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;**383**:166–75.
88 Errington TM, Iorns E, Gunn W, Tan FE, Lomax J, Nosek BA. An open investigation of the reproducibility of cancer biology research. *eLife* 2014;**3**:e04333.
89 Aird F, Kandela I, Mantis C. Replication Study: BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *eLife* 2017;**6**:e21253.
90 Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* 2011;**146**:904–17.
91 Brisco MJ, Latham S, Bartley PA, Morley AA. Incorporation of measurement of DNA integrity into qPCR assays. *Biotechniques* 2010;**49**:893–7.
92 Cone RW, Hobson AC, Huang ML. Coamplified positive control detects inhibition of polymerase chain reactions. *J Clin Microbiol* 1992;**30**:3185–9.
93 Nolan T, Hands RE, Ogunkolade BW, Bustin SA. SPUD: a qPCR assay for the detection of inhibitors in nucleic acid preparations. *Anal Biochem* 2006;**351**:308–10.
94 Huggett JF, Novak T, Garson JA, Green C, Morris-Jones SD, Miller RF *et al.* Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. *BMC Res Notes* 2008;**1**:70.
95 Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* 2006;**27**:126–39.
96 Auer H, Lyianarachchi S, Newsom D, Klisovic MI, Marcucci G, Kornacker K. Chipping away at the chip bias: RNA degradation in microarray analysis. *Nat Genet* 2003;**35**:292–3.

97 Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, Eveno E *et al.* Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res* 2005;**33**:e56.

98 Swift GH, Peyton MJ, MacDonald RJ. Assessment of RNA quality by semi-quantitative RT-PCR of multiple regions of a long ubiquitous mRNA. *Biotechniques* 2000;**28**:524–31.

99 Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 2006;**7**:3.

100 Brisco MJ, Morley AA. Quantification of RNA integrity and its use for measurement of transcript number. *Nucleic Acids Res* 2012;**40**: e144.

101 Bustin SA. *A-Z of Quantitative PCR*. La Jolla, CA: IUL Press; 2004.

102 Perez-Novo CA, Claeys C, Speleman F, Van Cauwenberge P, Bachert C, Vandesompele J. Impact of RNA quality on reference gene expression stability. *Biotechniques* 2005;**39**:52, 54, 56.

103 Vermeulen J, De Preter K, Lefever S, Nuytens J, De Vloed F, Derveaux S *et al.* Measurable impact of RNA quality on gene expression results from quantitative PCR. *Nucleic Acids Res* 2011;**39**: e63.

104 de Cremoux P, Valet F, Gentien D, Lehmann-Che J, Scott V, Tran-Perennou C *et al.* Importance of pre-analytical steps for transcriptome and RT-qPCR analyses in the context of the phase II randomised multicentre trial REMAGUS02 of neoadjuvant chemotherapy in breast cancer patients. *BMC Cancer* 2011;**11**:215.

105 Demeke T, Jenkins GR. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Anal Bioanal Chem* 2010;**396**:1977–90.

106 Hart C, Schochetman G, Spira T, Lifson A, Moore J, Galphin J *et al.* Direct detection of HIV RNA expression in seropositive subjects. *Lancet* 1988;**2**:596–9.

107 Delidow BC, Peluso JJ, White BA. Quantitative measurement of mRNAs by polymerase chain reaction. *Gene Anal Tech* 1989;**6**: 120–4.

108 Williams WV, Rosenbaum H, Weiner DB. Effect of RNA concentration on cDNA synthesis for DNA amplification. *PCR Methods Appl* 1992;**2**:86–8.

109 Pallansch L, Beswick H, Talian J, Zelenka P. Use of an RNA folding algorithm to choose regions for amplification by the polymerase chain reaction. *Anal Biochem* 1990;**185**:57–62.

110 Kuo KW, Leung MF, Leung WC. Intrinsic secondary structure of human TNFR-I mRNA influences the determination of gene expression by RT-PCR. *Mol Cell Biochem* 1997;**177**:1–6.

111 Brooks EM, Sheflin LG, Spaulding SW. Secondary structure in the 3′ UTR of EGF and the choice of reverse transcriptases affect the detection of message diversity by RT-PCR. *Biotechniques* 1995;**19**:814.

112 Todd J, Pachl C, White R, Yeghiazarian T, Johnson P, Taylor B *et al.* Performance characteristics for the quantitation of plasma HIV-1 RNA using branched DNA signal amplification technology. *J Acqui Immune Defic Syndr Hum Retrovirol* 1995;**10**(Suppl 2):S35–44.

113 Melo JV, Yan XH, Diamond J, Lin F, Cross NC, Goldman JM. Reverse transcription/polymerase chain reaction (RT/PCR) amplification of very small numbers of transcripts: the risk in misinterpreting negative results. *Leukemia* 1996;**10**:1217–21.

114 Chandler DP, Wagnon CA, Bolton HJ. Reverse transcriptase (RT) inhibition of PCR at low concentrations of template and its implications for quantitative RT-PCR. *Appl Environ Microbiol* 1998;**64**:669–77.

115 Zhang J, Byrne CD. Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. *Biochem J* 1999;**337**:231–41.

116 Keilholz U, Willhauck M, Scheibenbogen C, de Vries TJ, Burchill S. Polymerase chain reaction detection of circulating tumour cells. EORTC Melanoma Cooperative Group, Immunotherapy Subgroup. *Melanoma Res* 1997;**7**(Suppl 2):S133–41.

117 Keilholz U, Willhauck M, Rimoldi D, Brasseur F, Dummer W, Rass K *et al.* Reliability of reverse transcription-polymerase chain reaction (RT-PCR)-based assays for the detection of circulating tumour cells: a quality-assurance initiative of the EORTC Melanoma Cooperative Group. *Eur J Cancer* 1998;**34**:750–3.

118 Jung R, Ahmad-Nejad P, Wimmer M, Gerhard M, Wagener C, Neumaier M. Quality management and influential factors for the detection of single metastatic cancer cells by reverse transcriptase polymerase chain reaction. *Eur J Clin Chem Clin Biochem* 1997;**35**:3–10.

119 Freeman WM, Walker SJ, Vrana KE. Quantitative RT-PCR: pitfalls and potential. *Biotechniques* 1999;**26**:112–5.

120 Gibson UE, Heid CA, Williams PM. A novel method for real time quantitative RT-PCR. *Genome Res* 1996;**6**:995–1001.

121 Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* 2004;**15**:155–66.

122 Nam DK, Lee S, Zhou G, Cao X, Wang C, Clark T *et al.* Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proc Natl Acad Sci USA* 2002;**99**:6152–6.

123 Bustin SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* 2002;**29**:23–39.

124 Stahlberg A, Hakansson J, Xian X, Semb H, Kubista M. Properties of the reverse transcription reaction in mRNA quantification. *Clin Chem* 2004;**50**:509–15.

125 Stahlberg A, Kubista M, Pfaffl M. Comparison of reverse transcriptases in gene expression analysis. *Clin Chem* 2004;**50**:1678–80.

126 Bustin S, Dhillon HS, Kirvell S, Greenwood C, Parker M, Shipley GL *et al.* Variability of the reverse transcription step: practical implications. *Clin Chem* 2015;**61**:202–12.

127 Sanders R, Mason DJ, Foy CA, Huggett JF. Evaluation of digital PCR for absolute RNA quantification. *PLoS ONE* 2013;**8**:e75296.

128 Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonak J, Lind K *et al.* The real-time polymerase chain reaction. *Mol Aspects Med* 2006;**27**:95–125.

129 Lee TK, Murthy SR, Cawley NX, Dhanvantari S, Hewitt SM, Lou H *et al.* An N-terminal truncated carboxypeptidase E splice isoform induces tumor growth and is a biomarker for predicting future metastasis in human cancers. *J Clin Invest* 2011;**121**:880–92.

130 Torelli R, Sanguinetti M, Moody A, Pagano L, Caira M, De Carolis E *et al.* Diagnosis of invasive aspergillosis by a commercial real-time PCR assay for Aspergillus DNA in bronchoalveolar lavage fluid samples from high-risk patients compared to a galactomannan enzyme immunoassay. *J Clin Microbiol* 2011;**49**:4273–8.

131 Rossen L, Norskov P, Holmstrom K, Rasmussen OF. Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions. *Int J Food Microbiol* 1992;**17**:37–45.

132 Zheng Z, Cohn MJ. Developmental basis of sexually dimorphic digit ratios. *Proc Natl Acad Sci USA* 2011;**108**:16289–94.

133 Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010;**465**:1033–8.
134 Buzard GS, Baker D, Wolcott MJ, Norwood DA, Dauphin LA. Multi-platform comparison of ten commercial master mixes for probe-based real-time polymerase chain reaction detection of bioterrorism threat agents for surge preparedness. *Forensic Sci Int* 2012;**223**:292–7.
135 Picard-Meyer E, Peytavin de Garam C, Schereffer JL, Marchal C, Robardet E, Cliquet F. Cross-platform evaluation of commercial real-time SYBR green RT-PCR kits for sensitive and rapid detection of European bat Lyssavirus type 1. *Biomed Res Int* 2015;**2015**:839518.
136 Aird D, Ross MG, Chen WS, Danielsson M, Fennell T, Russ C *et al.* Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* 2011;**12**:R18.
137 Dabney J, Meyer M. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *Biotechniques* 2012;**52**:87–94.
138 Pan W, Byrne-Steele M, Wang C, Lu S, Clemmons S, Zahorchak RJ *et al.* DNA polymerase preference determines PCR priming efficiency. *BMC Biotechnol* 2014;**14**:10.
139 Wacker MJ, Godard MP. Analysis of one-step and two-step real-time RT-PCR using SuperScript III. *J Biomol Tech* 2005;**16**:266–71.
140 Chowdhury P, Khan SA, Dutta P, Topno R, Mahanta J. Characterization of West Nile virus (WNV) isolates from Assam, India: insights into the circulating WNV in northeastern India. *Comp Immunol Microbiol Infect Dis* 2014;**37**:39–47.
141 D'Aquila RT, Bechtel LJ, Videler JA, Eron JJ, Gorczyca P, Kaplan JC. Maximizing sensitivity and specificity of PCR by pre-amplification heating. *Nucleic Acids Res* 1991;**19**:3749.
142 Scalice ER, Sharkey DJ, Daiss JL. Monoclonal antibodies prepared against the DNA polymerase from Thermus aquaticus are potent inhibitors of enzyme activity. *J Immunol Methods* 1994;**172**:147–63.
143 Sharkey DJ, Scalice ER, Christy KGJ, Atwood SM, Daiss JL. Antibodies as thermolabile switches: high temperature triggering for the polymerase chain reaction. *Biotechnology (N Y)* 1994;**12**:506–9.
144 Stevens AJ, Appleby S, Kennedy MA. Many commercial hot-start polymerases demonstrate activity prior to thermal activation. *Biotechniques* 2016;**61**:293–6.
145 Burns MJ, Nixon GJ, Foy CA, Harris N. Standardisation of data from real-time quantitative PCR methods - evaluation of outliers and comparison of calibration curves. *BMC Biotechnol* 2005;**5**:31.
146 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;**25**:402–8.
147 Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;**29**:E45.
148 Bustin SA. Real-time, fluorescence-based quantitative PCR: a snapshot of current procedures and preferences. *Expert Rev Mol Diagn* 2005;**5**:493–8.
149 Bustin S, Bergkvist A, Nolan T. In silico tools for qPCR assay design and data analysis. *Methods Mol Biol* 2011;**760**:283–306.
150 Jin N, He K, Liu L. qPCR-DAMS: a database tool to analyze, manage, and store both relative and absolute quantitative real-time PCR data. *Physiol Genomics* 2006;**25**:525–7.
151 O'Connor W, Runquist EA. Error minimization algorithm for comparative quantitative PCR analysis: Q-Anal. *Anal Biochem* 2008;**378**:96–8.
152 Ritz C, Spiess AN. qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* 2008;**24**:1549–51.
153 Pabinger S, Thallinger GG, Snajder R, Eichhorn H, Rader R, Trajanoski Z. QPCR: Application for real-time PCR data management and analysis. *BMC Bioinformatics* 2009;**10**:268.
154 Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 2007;**8**:R19.
155 Karlen Y, McNair A, Perseguers S, Mazza C, Mermod N. Statistical significance of quantitative PCR. *BMC Bioinformatics* 2007;**8**:131.
156 Svec D, Tichopad A, Novosadova V, Pfaffl MW, Kubista M. How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. *Biomol Detect Quantif* 2015;**3**:9–16.
157 Pabinger S, Rödiger S, Kriegner A, Vierlinger K, Weinhäusel A. A survey of tools for the analysis of quantitative PCR (qPCR) data. *Biomol Detect Quantif* 2014;**1**:23–33.
158 Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JP. Public availability of published research data in high-impact journals. *PLoS ONE* 2011;**6**:e24357.
159 Stolovitzky G, Cecchi G. Efficiency of DNA replication in the polymerase chain reaction. *Proc Natl Acad Sci USA* 1996;**93**:12947–52.
160 Peirson SN, Butler JN, Foster RG. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Res* 2003;**31**:e73.
161 Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ *et al.* Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 2009;**37**:e45.
162 Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* 2003;**339**:62–6.
163 Rutledge RG. Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications. *Nucleic Acids Res* 2004;**32**:e178.
164 Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M *et al.* Primer sequence disclosure: a clarification of the MIQE guidelines. *Clin Chem* 2011;**57**:919–21.
165 Bustin SA. Why the need for qPCR publication guidelines?–The case for MIQE. *Methods* 2010;**50**:217–26.
166 Pajer K, Andrus BM, Gardner W, Lourie A, Strange B, Campo J *et al.* Discovery of blood transcriptomic markers for depression in animal models and pilot validation in subjects with early-onset major depression. *Transl Psychiatry* 2012;**2**:e101.
167 Redei EE, Andrus BM, Kwasny MJ, Seok J, Cai X, Ho J *et al.* Blood transcriptomic biomarkers in adult primary care patients with major depressive disorder undergoing cognitive behavioral therapy. *Transl Psychiatry* 2014;**4**:e442.

