The Impact of data selection and fitting on SAR estimation for magnetic nanoparticle heating

Hattie L. Ring1,*, Anirudh Sharma2,*, Robert Ivkov3,†, John C. Bischof2,4,†
1Center for Magnetic Resonance Imaging, University of Minnesota, Minneapolis, USA
2Department of Mechanical Engineering, University of Minnesota, Minneapolis, USA
3Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University School of Medicine, Baltimore, USA
4Department of Biomedical Engineering, University of Minnesota, Minneapolis, USA

Abstract

Magnetic fluid heating has great potential in the fields of thermal medicine and cryopreservation. However, variations among experimental parameters, analysis methods and experimental uncertainty make quantitative comparisons of results among laboratories difficult. Herein, we focus on the impact of calculating the specific absorption rate (SAR) using Time-Rise and Box-Lucas fitting. Time-Rise assumes adiabatic conditions, which is experimentally unachievable, but can be reasonably assumed (quasi-adiabatic) only for specific and limited evaluation times when heat loss is negligible compared to measured heating rate. Box-Lucas, on the other hand, accounts for heat losses but requires longer heating. Through retrospective analysis of data obtained from two laboratories, we demonstrate measurement time is a critical parameter to consider when calculating SAR. Volumetric SAR were calculated using the two methods and compared across multiple iron-oxide nanoparticles. We observed the lowest volumetric SAR variation from both fitting methods between 1 – 10 W/mL, indicating an ideal SAR range for heating measurements. Furthermore, our analysis demonstrates that poorly chosen fitting method can generate reproducible but inaccurate SAR. We provide recommendations to select measurement time for data analysis with either Modified Time-Rise or Box-Lucas method, and suggestions to enhance experimental precision and accuracy when conducting heating experiments.

Corresponding author contact information: John Bischof, Department of Mechanical Engineering, University of Minnesota, 111 Church Street SE, Minneapolis, MN 55455, USA. bischof@umn.edu.
* designates equal contribution by Primary co-authors
† designates equal contribution by Senior co-authors

Declaration of interests
R.I. is an inventor on nanoparticle patents, and all patents are assigned to either The Johns Hopkins University or Aduro BioTech, Inc. R.I. also consults for Imagion Biosystems, a company developing imaging with magnetic iron oxide nanoparticles. R.I. received partial funding from the National Cancer Institute (5R01CA194574–02 and 5R01CA247290). J.C.B. is an inventor on several patents that use iron oxide and gold nanoparticles to heat biomaterials for regenerative medicine purposes. These patents are assigned to the University of Minnesota. H.R., A.S. and J.C.B. received partial funding from NIH R01 HL135046 & R01 DK117425 which are related to iron oxide nanoparticle rewarming of vitrified biomaterials. J.C.B. was also supported by the Bakken Chair for Engineering in Medicine at the University of Minnesota. All other authors report no conflicts of interest. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official view of the Johns Hopkins University, University of Minnesota, NIH, or other funding agencies.
Keywords
Magnetic iron oxide nanoparticles; magnetic fluid hyperthermia; data analysis; nanowarming; specific absorption rate calculation

Introduction
Magnetic nanoparticle or magnetic fluid hyperthermia (MFH) is a cancer treatment indicated as an adjunct treatment for recurrent glioblastoma with radiation therapy [1, 2]. It comprises direct delivery of a magnetic fluid, i.e. aqueous suspension of iron oxide nanoparticles, into the tumor that generates heat when exposed to an alternating magnetic field [1, 3–6]. The therapeutic agent is heat produced by the nanoparticles, and thus potency or heating capability of the suspension is a critical parameter for treatment planning and execution [7–9]. Additionally, emerging clinical applications, such as neuromodulation and cryopreservation of tissues, exploit controlled heat production with magnetic nanoparticles [10–12]. A direct comparison of iron oxide nanoparticle (IONP) heating capability for MFH across laboratories has proven unreliable, thus challenging quantitative comparisons among IONP formulations to determine clinical suitability, and inhibiting development of universal standardized methodologies for clinical translation [13, 14].

Hysteresis loss power heat generated by magnetic nanoparticles is often measured with simple constant pressure (atmospheric) calorimetric methods, from which normalized heating power is estimated. The calorimeter often comprises a thermally insulated vessel with thermometer that is placed within the induction coil. A known volume of sample, or nanoparticle suspension having a known concentration, is placed into the vessel. Heat generated by the nanoparticles with AMF is absorbed by the suspending medium (e.g. water), causing a temperature change (ΔT) which is measured. It is typically assumed that the temperature of the sample is reasonably homogeneous, and that the temperature probe measurement accurately reflects the temperature throughout the sample volume. Using the sample specific heat (i.e. typically assumed to be the medium), estimates of the energy absorbed are then performed from analysis of changes of temperature within a specified time interval. An inherent assumption with this methodology is that the temperature changes measured in the sample occur only from heat generated by the nanoparticles in response to the AMF. Heat transfer between the sample and the environment, if unaccounted for, will inevitably generate erroneous values of loss power heating. The method of analysis, therefore must be carefully chosen to incorporate the limitations of the measurement methods and/or apparatus, and the data collected. It is often the case that estimation of heating power implicitly assumes adiabatic conditions (Eq. 1). This condition may not be strictly possible given that the induction coil can be either heat source or heat sink, depending on specific properties of its operation (e.g. current load, voltage, cooling capacity of coolant, etc.). Further, this method is limited by errors within temperature measurements, which arise from limitations of equipment resolution, experimental setup (heat exchange with surroundings, field inhomogeneities/instabilities, chemical/morphological changes in the sample over time), execution, uncertainty in iron mass determination and thermometry measurement [15–18]. Additionally, various methods are used to fit and analyse data, further
contributing to reduced consistency and reliability [8, 18, 19]. Often method(s) used to estimate normalized dissipated power from temperature-time measurements is/are fixed without considering the underlying assumptions for the calculations, and limitations of the apparatus used to measure heat generated. For iron containing magnetic nanoparticles, specific loss power (SLP) is often reported as a function of iron mass to indicate IONP heating capability (W/g Fe) [3, 14]. Assuming no heat exchange with surroundings, no work performed, no chemical or physical (change of state) changes in the sample, and constant pressure, the relationship between SLP, power ($P$) and the change in temperature are:

$$ SLP = \frac{P}{m_{Fe}} = \frac{m_s \cdot c \cdot \frac{dT}{dt}}{m_{Fe}} $$

(1)

where $c$ is the specific heat (J/ g°C), $m_s$ = sample mass, $m_{Fe}$ = the mass of iron (g), and $dT/dt$ is the temperature change within a defined time interval, or first derivative of the temperature as a function of time (°C/s). Volumetric SAR can be described as power output normalized to sample volume (SAR$_v$, W/mL):

$$ SAR_v = \rho \cdot c \cdot \frac{dT}{dt} = SLP \cdot c_{Fe} $$

(2)

where, $\rho$ is the sample density, $c_{Fe}$ is the iron concentration in the sample. While SAR$_v$ and SLP can be easily converted as shown in Equation (2), there are advantages to working with SAR$_v$ alone from the point of view of calorimetry and heat transfer. First, the sample density and specific heat (i.e. water, tissue or other) are largely unchanged for low concentrations of nanoparticles and can be directly measured prior to calorimetry. This means that SAR$_v$ can be reasonably estimated based on the $dT/dt$ even without precise knowledge of iron content or SLP as long as our calorimetric assumptions hold. This means that estimation of error in fitting SAR$_v$, a main feature of the present work, is more directly correlated to the calorimetry itself and not compounded by possible errors from iron concentration estimation [21, 22, 34, 35]. In addition to being a concept that can be directly related to the calorimetric measurement it also becomes an estimate of heat that can be produced by iron oxide nanoparticles in complex systems (i.e. tissues/organs) for known IONP concentrations [8, 20–22]. This in turn can be used in computational heat transfer modelling efforts employing the bioheat transfer equation. We report results of comparisons between fitting methods and power output from heating measurements obtained for several samples in two different laboratories. For the evaluated datasets, sample density was consistent; however, IONP concentrations among samples varied. Therefore, we use SAR$_v$ for comparison because it correlates with power for the scope of the evaluated datasets.

Using the Time-Rise method, SLP, and SAR$_v$ are calculated from the slope $(dT/dt)$ by fitting $T$ vs $t$ [18,19]. The Time-Rise calculation can be used during low heat loss conditions (i.e., a well-insulated system, short experimental evaluation times, and low SAR$_v$).

Box-Lucas analysis assumes a non-adiabatic system where the convective heat losses can be approximated as a loss term with the difference between the sample and ambient temperature. Wildeboer et. al. provide a derivation of the Box-Lucas equation based on this
assumption which results in an exponential growth dependence of temperature as a function of time [14] (Figure 1):

\[ T = \frac{P}{L} \left(1 - e^{-\frac{t}{cL}}\right) + T_0 \]  

where, \( L \) is the heat loss from the sample, and \( P \) is the power generated by the sample. Multiplying the amplitude \( (PL) \) and exponential factor \( (L/c) \) from fitting the Box-Lucas equation (Eq. 4) provides \( P/c \) is shown in Eqs. 1 & 2 to be equivalent to \( dT/dt \) and can be used to calculate SLP, and SAR\(_v\). We note that radiative heat losses can be linear and important when \( \Delta T \ll T \).

Several alternate calculation methods have been proposed to compensate for heat losses. Specifically, measurements of the temperature steady state occurring at the end of the heating phase of the experiment or the temperature decay during the cooling phase can be used to directly calculate the heating losses of the system. (Figure 1) [14–16, 23, 24]. Calculation of heat losses allows for the determination if linear heat losses are dominant [14]. For convenience, SAR\(_v\) is often calculated from the heating phase of the experiment using either Time-Rise (i.e. initial slope), or Box-Lucas fitting. Due to the different assumptions, discrepancies between estimates of SAR\(_v\) obtained from Time-Rise and Box-Lucas fitting can occur [18, 19, 25]. The many sources of error within the measuring equipment and physical sample can affect the accuracy and precision of SAR\(_v\) [14, 16, 24–26]. The selection of Time-Rise or Box-Lucas fitting to calculate SAR\(_v\) will variably affect precision and accuracy depending on the selected measurement time interval and variance within the data. Herein, we perform a direct comparison between Box-Lucas and Time-Rise fitting methods by reanalysing previously published datasets to determine the appropriate conditions (optimal SAR\(_v\) range, length of evaluation time and starting times) for using Box-Lucas and Time-Rise fitting methods.

**Materials and Methods**

**Datasets:**

These two datasets were collected from several different IONPs, under different conditions (frequency, magnetic field strength, IONP concentration), from two different laboratories (Johns Hopkins University & University of Minnesota). The acquisition time of the temperature-time datasets was sufficient to enable a comparison of variations in evaluation time.

**Retrospective Data Analysis:**

Retrospective data analysis was performed with MATLAB 2012b on temperature-time datasets published in Bordelon et. al. and Etheridge et. al.[19, 23]. Four different evaluation time scenarios were compared: (1) 30s; (2) 60s; (3) entire dataset (varied); (4) \( R^2 \) optimization. The \( R^2 \) optimization is similar to the Pearson correlation coefficient \( (R^2) \) analysis used in Etheridge et. al., however, instead of the use of conditional statements, the \( R^2 \) optimization used here was simplified to select only the evaluation time based on the highest \( R^2 \) value for both Time-Rise and Box-Lucas fitting [27]. We compared the effects of...
(1) heating start time; (2) 5s after heating start time; (3) following the criteria for the quasi-adiabatic regime. The discrete derivative required to satisfy the quasi-adiabatic criteria, was observed to amplify noise within the measurement. Within our assessment, the identification of the quasi-adiabatic criteria was tested across a 10s timeframe and was met when the slope of the time change and the average of the first derivative were within 5%. Noise was removed from the first derivative using total-variation regularization [28].

Once the timeframe for evaluation was defined, Box-Lucas and Time-Rise fitting were performed, and the SAR$_v$ and root mean standard error (RMSE) were calculated. Within the Bordelon et. al. dataset, precision of data was assessed based on the RMSE of the fitted data. The temperature-time datasets from Etheridge et. al. enabled measurement precision to be evaluated based on the standard deviation of measured replicates. Variations in calculated values for SAR$_v$ were reported as an indication of measurement accuracy.

### Results and Discussion:

#### Retrospective Analysis:

Bordelon et. al. and Etheridge et. al.[19, 23] demonstrate two similar experiments with different approaches for calculating SAR$_v$. Bordelon et. al. [19] reported the amplitude dependent SLP of three commercially available IONPs (BNF-starch, nanomag-D-spio, and Feridex). Etheridge et. al. compared two commercially available IONPs (starch-coated Micromod and Ferrotec EMG-308). The experimental setup described within both experiments used solenoid coils and the IONPs were suspended in water.

A comparison among the temperature-time datasets measured the peak-to-peak noise in Etheridge et. al. (0.13 °C) was over twice the peak-to-peak noise in Bordelon et. al. (0.06 °C). This comparison of noise is indirect, because Etheridge et. al. measured heating during 30s of negligible temperature change with the coil on in control samples, while Bordelon et. al. was measured during 30s prior to turning on power to the inductor in addition to measuring water blank control samples. However, the anticipated noise within the temperature measurement with the coil on is not anticipated to double the peak-to-peak noise. As an example of minor differences in experimental setup is the handling of inhomogeneous fields from the solenoid coil. Bordelon et. al. used a well characterized four-turn coil with a measured homogeneous field region to determine sample placement [19, 29]. Etheridge et. al. simulated the inhomogeneity of their 2.75-turn coil and adjusted the reported magnetic field amplitude.

Bordelon et. al. performed heating measurements at a large range of magnetic field amplitudes (4 – 94 kA/m) resulting in a broad range of reported SLP (0.26 – 537 W/g Fe). The reported SLP for each condition was calculated using both the Time-Rise and Box-Lucas fit methods. The evaluation times varied from 7 – 30 s for Time-Rise fittings and 6 – 96 s for Box-Lucas fittings. The variation in time frame of the measurements was caused by stopping the heating experiment at 70 °C to avoid boiling the aqueous solution [19].

Etheridge et. al. varied frequency (190 – 370 kHz), magnetic field strength (5 – 20 kA/m) and IONP concentration (0.05 –1 mg Fe/mL). These variations produced a smaller range of
reported SLP values (20 – 500 W/g Fe) than Bordelon et. al.; however, Etheridge et. al. performed six replicates for each measured condition allowing for the assessment of reproducibility. The evaluation times varied from 10 to 60 s and the start time ranged from 0 – 10 s after heating began. The evaluation time was shorter than the total heating time, 300s. Only Time-Rise calculations were performed and the analysis timeframe for these datasets was selected based on a series of conditional criteria to minimize the fit error assessed by the Pearson correlation coefficient ($R^2$) [27].

Soetaert et. al. demonstrated the impact of analysis time frame on the calculated SLP [18]. The effect of evaluation time is intuitive based on the assumptions made for Time-Rise and Box-Lucas fitting and is discussed in further detail below. The selection of start time affects the calculation of $SAR_v$, due to the observed non-linear response between starting the heating experiment and measurable changes in temperature. The cause of the non-linear response has been attributed to the response of inductive heating mechanisms, diffusion of heat within the sample, temperature probe delay, thermal mixing, colloidal dispersion of IONPs, field homogeneity, and heat transfer properties of the sample [14, 15, 18, 24, 30–32]. Regardless of the cause, the non-linear response at the beginning of the heating experiment needs to be adjusted and reported. Soetaert et. al. [18] demonstrated a method to select the dataset within the ‘quasi-adiabatic’ regime (i.e. the time frame where temperature changes are approximately linear within a non-adiabatic system). The quasi-adiabatic regime is defined by comparing $dT/dt$ calculated from the first derivative and the slope of the dataset [18].

Case 1: Bordelon et. al.[19]

The range of the $SAR_v$ and RMSE calculated with all start and evaluation time scenarios are shown in Figure 2, demonstrating the most stable results despite fitting method, for a $SAR_v$ between 1 and 10 W/mL. At low $SAR_v$ (<1 W/mL) an increase in RMSE and $SAR_v$ variation is observed, indicating the lower limit of $SAR_v$ measurement (see red circle, Figure 2). The comparison of fitting parameters was challenging when $SAR_v > 1.2$ W/mL, due to the short duration (< 30s) removing the ability to evaluate the different evaluation times (30s, 60s, and full data range). The shortened timeframe results in a wider range in $SAR_v$ for Box-Lucas than Time-Rise. The $SAR_v$ range from 0.2 – 1.2 W/mL was selected to compare multiple evaluation times and adjust the starting point, because more data allowed for a comparison among all evaluation times. $SAR_v$ did not show high variability (Supplemental Figures S1 and S2).

Figures 3.A and 3.B shows the impact of evaluation time on the variation of $SAR_v$ and RMSE. Within this range of $SAR_v$ values, it is observed that the Time-Rise fitting method is more sensitive than the Box-Lucas fitting method to the chosen evaluation time. This is expected since with increasing evaluation time the effect of heat loss on the dataset becomes noticeable and the assumed adiabatic assumption needed for Time-Rise is violated. More specifically within the dataset (Figures S1 & S2), we observed that the calculated $SAR_v$ decreased while RMSE increased with increasing evaluation time. Therefore, $SAR_v$ is underestimated when a long evaluation time which includes heat losses is selected. RMSE is
more sensitive to heat losses caused by long evaluation times and can be used to select an appropriate evaluation time.

Figures 3.C and 3.D show Box-Lucas fitting to be more sensitive to start time than Time-Rise fitting. While not observed within this comparison, if the selection of the start time creates a dataset that includes the initial heating non-linear response, the SAR_v will be underestimated with Time-Rise fitting due to the decrease in the slope of the linear fit [14]. The Box-Lucas fitting method shows an increase in SAR_v and a decrease in RMSE (Figures S1 & S2) caused by selecting a start time that does not include the initial heating non-linear response (i.e. analysis is performed 5 s after the experimental start time). Therefore, when the timeframe selected incorporates data from the initial 5 s, inaccurate results are acquired. For Box-Lucas fitting, including the initial non-linear heating response causes a decrease in the exponential factor to fit the curvature of the temperature change as a function of time. For both Time-Rise and Box-Lucas fitting methods, the assumption is that heat is being deposited into the sample/system; the initial heating non-linear response indicates a timeframe where steady-state heat generation is not observed and should not be used for analysis. It is important to note the variation in SAR_v is relatively low (< 10%) with changes in the evaluation time and start time; however, the effects of selecting the time frame of the dataset can be observed with the RMSE.

Finally, rigidly selected evaluation and start times were compared with automated selection methods for the evaluation time ($R^2$ optimization) and start time (quasi-adiabatic criteria). Both automated methods resulted in a reduction in RMSE (Figure S2). For the Time-Rise fitting method, the calculated SAR_v increased with increasing heat production when using $R^2$ optimization. The $R^2$ optimization was able to reduce the evaluation time to as short as 5 s, which allowed the dataset to comply with the adiabatic assumption of Time-Rise as experimental conditions (i.e. IONP, frequency, and magnetic field amplitudes) resulted in an increase of SAR_v. Therefore, the observed increase in SAR_v with a decrease in RMSE indicates a shortened dataset removing heat loss effects. $R^2$ optimization had negligible impact on SAR_v or RMSE calculated through the Box-Lucas fitting method. Furthermore, the use of the quasi-adiabatic selection for start time, did not have a significant effect on SAR_v or RMSE compared to the 5 s delay start time.

**Case 2: Etheridge et. al.[27]**—The range of the SAR_v and RMSE calculated with all start and evaluation time scenarios are in Supplemental Figure S3. The Etheridge et. al. dataset demonstrated higher RMSE compared to the Bordelon et al., likely caused by the higher noise within the experimental setup. Similar to observations with Bordelon et. al., higher variations in SAR_v based on fitting parameters were evident as SAR_v decreased, Time-Rise fitting results are more sensitive to evaluation time, and SAR_v calculated with Box-Lucas fitting is more sensitive to changes in start time. Evaluation of the full dataset time-interval caused Time-Rise SAR_v to be approximately 30% higher than the average SAR_v (Supplemental Figure S4) and RMSE to be over 350% higher than SAR_v (Supplemental Figure S5). The large effect on Time-Rise SAR_v and RMSE is caused by the 300 s evaluation time causing a violation of the adiabatic assumption.
Very few datasets were observed to have RMSE < 10% SAR\textsubscript{v} (Supplemental Figure S5); however, the standard deviation of replicate measurements was < 10% SAR\textsubscript{v} with rigid time-frame selection (Figure 4). Although a very high RMSE (> 500% SAR\textsubscript{v}) was calculated for many datasets, this did not impact the repeatability of using the Time-Rise fitting for analysis. It is important to recognize that a high RMSE indicates a deviation from the assumptions of the fitting method. For Time-Rise fitting, as the SAR\textsubscript{v} increases the heating losses become sufficient to violate the adiabatic assumption. Similarly, a violation of the linear heating losses assumption for Box-Lucas fitting should occur as SAR\textsubscript{v} increases, although this was not observed for the range of SAR\textsubscript{v} evaluated. Therefore, it is important to report RMSE to indicate the accuracy of the assumptions made with the fitting method used to calculate SAR\textsubscript{v} and to report the standard deviation as an indication of the ability to repeat measurements.

Figure 4 demonstrates the impact evaluation time has the on repeatability of measurement, as well as the impact of using an automated method, such as the \textit{R}^2 optimization for selecting the evaluation time. The selection of start time, resulted in the variations of the calculated SAR\textsubscript{v}, however, no trend based on the start time selected is visible (Supplemental Figure S6). Furthermore, automating the start point selection using the quasi-adiabatic method did not result in an increased variation in the repeatability of the measurements. Within the datasets, five data points were removed as outliers using the Q-test for comparing SAR\textsubscript{v} across the different fitting parameters. These outliers only occurred for automated fittings (both \textit{R}^2 optimization and quasi-adiabatic selection). Additionally, the high noise within the discrete derivative was problematic and was removed by using a total-variation regularization [28]. The increase in this effect as SAR\textsubscript{v} decreases indicates automated methods for selecting the dataset timeframe become less robust with increasing noise.

Table 1 summarizes recommendations for parameters to report for fitting with either linear-rise or Box-Lucas, depending on the range of SAR\textsubscript{v}. Overall, the evaluation of RMSE, for a dataset and a fitting method (Time-Rise vs Box-Lucas) with its underlying assumptions, will help guide the selection of evaluation time and start time for that fitting method. Generally, while employing the Time-Rise fitting method, short time intervals (<30 s) that satisfy the quasi-adiabatic criteria are suitable for SAR\textsubscript{v} estimation (small RMSE). Box-Lucas estimation, on the other hand, is less influenced by the length of the evaluation time interval, if losses in the system are linear. However, high SAR\textsubscript{v} values (e.g., SAR\textsubscript{v} >10 W/mL) can result in high RMSE using the Box-Lucas method because of non-linear losses in the measurement. In such cases, Time-Rise fit, constrained by the quasi-adiabatic criteria, can be a better fitting method. For both fitting methods, a target SAR\textsubscript{v} range of 1–10 W/mL (specific to the datasets evaluated here), appears to generate consistent SAR\textsubscript{v} values. An initial calibration of c(Fe) vs SAR\textsubscript{v} across a broad range of iron concentration (e.g. 1–10 mg Fe/mL) may be useful to ascertain an optimal SAR\textsubscript{v} range for a specific laboratory and setup. For both fitting methods, it is recommended to select a start time from the T vs t data that does not include the initial heating non-linear response (i.e. remove temperature data in the 0–5 s from the onset of heating in the fitting analysis). Additionally, reduction in peak-to-peak noise in the temperature measurement to ≤0.06 °C is desirable to reduce its contribution to the RMSE. Calibration of the AMF coil, calorimeter, temperature probes and...
the use of appropriate insulation are essential to characterizing and reducing the peak-to-peak noise.

The challenges of reliably comparing SAR\textsubscript{v} across laboratories are related to the complex nature of the heat losses involved with the heating system and the sample. Most research findings report field strength, field amplitude, and sample information; however, many details related to the experimental setup are necessary to understand the heat losses within the system. Calibration of calorimeter, field properties, and other components of the apparatus are rarely performed, presenting significant challenges for comparisons absent standard reference materials. Additionally, the precision and accuracy with which the iron content is quantified will affect the calculated SLP. While ICP-MS is the gold-standard for iron quantification due to its sensitivity (ng/L detection limits), less expensive spectrophotometric assays have been optimized to measure iron content in 0.1–1μg/mL range [34,35], which should be suitable for SLP evaluation. Finally, the sample volume used for measurements has been shown to critically affect the results, likely due to heat transfer between sample and environment [25,26]. We recommend researchers report at least the experimental details listed in Table 2 to indicate the sample preparation and heat losses within an experimental setup. Finally, it cannot be overlooked that each experimental setup has an optimal range of SAR\textsubscript{v} and adjustment of IONP concentration to be within that range can reduce variation results caused by fitting method.

Conclusion

Valid selection between Time-Rise and Box-Lucas analysis for heating experiments depends on the experimental setup and the systematic error introduced in the calorimetric measurement. If a heating range appropriate for the apparatus is selected, where variation in reported SAR\textsubscript{v} is minimal, the choice between parameters for analysis are within 10% of each other. For the experimental measurements evaluated here, this range was 1 – 10 W/mL. Recommendations for selecting between Time-Rise or Box-Lucas parameters are provided in Table 1.

Outside of the optimal SAR\textsubscript{v} measurement range, a consistent analysis method can allow for repeatable SAR\textsubscript{v} results, although accuracy is questionable. Selection of analysis methods and parameters can hardly compensate for poor quality data obtained. Proper calibration of calorimeter, thermometers, AMF field and frequency, etc. are absolute requirements for conducting such measurements. The assumptions of Time-Rise and Box-Lucas fitting should be considered when selecting evaluation time and start time, although the underlying cause of this is presently unclear. A decrease in RMSE can act as a helpful metric for comparing the selection of evaluation time or start time; however, automated methods optimized on minimizing RMSE can produce higher variations in results if a measurement is noisy. In short, these anomalies should be considered as indications that systematic (or systemic) errors exist in the measurement methods or apparatus, bearing a re-examination.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.
Acknowledgements

We thank Michael Etheridge for conversations related to the development of this manuscript.

References:

1. Mahmoudi K, et al., Magnetic hyperthermia therapy for the treatment of glioblastoma: a review of the therapy’s history, efficacy and application in humans. International Journal of Hyperthermia, 2018: p. 1–13.
2. Ring HL, Bischof JC, and Garwood M, The Use and Safety of Iron-Oxide Nanoparticles in MRI and MFH, in Handbook - RF Safety, Shrivastava D, Editor. 2019, eMagRes.
3. Rosensweig RE, Heating magnetic fluid with alternating magnetic field. Journal of Magnetism and Magnetic Materials, 2002 252(0): p. 370–374.
4. Dennis CL, et al., Nearly complete regression of tumors via collective behavior of magnetic nanoparticles in hyperthermia. Nanotechnology, 2009 20(39): p. 395103.
5. Giustini AJ, Ivkov R, and Hoopes PJ, Magnetic nanoparticle biodistribution following intratumoral administration. Nanotechnology, 2011 22(34): p. 345101–345101.
6. Hoopes PJ, et al. Intratumoral Iron Oxide Nanoparticle Hyperthermia and Radiation Cancer Treatment. in SPIE Proceedings. 2007.
7. Southern P and Pankhurst QA, Commentary on the clinical and preclinical dosage limits of interstitially administered magnetic fluids for therapeutic hyperthermia based on current practice and efficacy models. International Journal of Hyperthermia, 2018 34(6): p. 671–686. [PubMed: 29046072]
8. Zhang J, et al., Quantification and biodistribution of iron oxide nanoparticles in the primary clearance organs of mice using T1 contrast for heating. Magnetic Resonance in Medicine, 2017 78(2): p. 702–712. [PubMed: 27667655]
9. Hoopes PJ, et al., In Vivo Imaging and Quantification of Iron Oxide Nanoparticle Uptake and Biodistribution. Proceedings of SPIE, 2012 8317: p. 83170R.
10. Chen R, et al., Wireless magnetothermal deep brain stimulation. Science, 2015 347(6229): p. 1477. [PubMed: 25765068]
11. Manuchehrabadi N, et al., Improved tissue cryopreservation using inductive heating of magnetic nanoparticles. Science Translational Medicine, 2017 9(379).
12. Sharma A, Bischof JC, and Finger EB, Liver Cryopreservation for Regenerative Medicine Applications. Regenerative Engineering and Translational Medicine, 2019.
13. Tong S, et al., Size-Dependent Heating of Magnetic Iron Oxide Nanoparticles. ACS Nano, 2017.
14. Wildeboer RR, Southern P, and Pankhurst QA, On the reliable measurement of specific absorption rates and intrinsic loss parameters in magnetic hyperthermia materials. Journal of Physics D: Applied Physics, 2014 47(49): p. 495003.
15. Lahiri BB, Ranoo S, and Philip J, Uncertainties in the estimation of specific absorption rate during radiofrequency alternating magnetic field induced non-adiabatic heating of ferrofluids. Journal of Physics D: Applied Physics, 2017 50(45): p. 455005.
16. Makridis A, et al., A standardisation protocol for accurate evaluation of specific loss power in magnetic hyperthermia. Journal of Physics D: Applied Physics, 2019 52(25): p. 255001.
17. Skumiel A, et al., Uses and limitation of different thermometers for measuring heating efficiency of magnetic fluids. Applied Thermal Engineering, 2016 100(Supplement C): p. 1308–1318.
18. Soetaert F, et al., Experimental estimation and analysis of variance of the measured loss power of magnetic nanoparticles. Scientific Reports, 2017 7(1): p. 6661. [PubMed: 28751720]
19. Bordelon DE, et al., Magnetic nanoparticle heating efficiency reveals magneto-structural differences when characterized with wide ranging and high amplitude alternating magnetic fields. Journal of Applied Physics, 2011 109(12): p. 124904.
20. Etheridge ML, et al., Accounting for biological aggregation in heating and imaging of magnetic nanoparticles. Technology, 2014 02(03): p. 214–228.
21. Hurley KR, et al., Predictable Heating and Positive MRI Contrast from a Mesoporous Silica-Coated Iron Oxide Nanoparticle. Mol. Pharm, 2016 13(7): p. 2172–83. [PubMed: 26991550]

22. Zhang J, et al., Quantifying iron-oxide nanoparticles at high concentration based on longitudinal relaxation using a three-dimensional SWIFT look-locker sequence. Magnetic Resonance in Medicine, 2014 71(6): p. 1982–1988. [PubMed: 24664527]

23. Teran FJ, et al., Accurate determination of the specific absorption rate in superparamagnetic nanoparticles under non-adiabatic conditions. Applied Physics Letters, 2012 101(6): p. 062413.

24. Landi GT, Simple models for the heating curve in magnetic hyperthermia experiments. Journal of magnetism and magnetic materials, 2013 326: p. 14–21.

25. Huang S, et al., On the measurement technique for specific absorption rate of nanoparticles in an alternating electromagnetic field. Measurement Science and Technology, 2012 23(3): p. 035701.

26. Wang SY, Huang S, and Borca-Tasciuc DA, Potential Sources of Errors in Measuring and Evaluating the Specific Loss Power of Magnetic Nanoparticles in an Alternating Magnetic Field. IEEE Transactions on Magnetics, 2013 49(1): p. 255–262.

27. Etheridge ML and Bischof JC, Optimizing Magnetic Nanoparticle Based Thermal Therapies Within the Physical Limits of Heating. Annals of Biomedical Engineering, 2013 41(1): p. 78–88. [PubMed: 22855120]

28. Chartrand R, Numerical Differentiation of Noisy, Nonsmooth Data. ISRN Applied Mathematics, 2011 2011: p. 11.

29. Bordelon DE, et al., Modified solenoid coil that efficiently produces high amplitude AC magnetic fields with enhanced uniformity for biomedical applications. Magnetics, IEEE Transactions on, 2012 48(1): p. 47–52.

30. Conde-Leborán I, Serantes D, and Baldomir D, Orientation of the magnetization easy axes of interacting nanoparticles: Influence on the hyperthermia properties. Journal of Magnetism and Magnetic Materials, 2015 380: p. 321–324.

31. Munoz-Menendez C, et al., The role of size polydispersity in magnetic fluid hyperthermia: average vs. local infra/over-heating effects. Physical Chemistry Chemical Physics, 2015 17(41): p. 27812–27820. [PubMed: 26437746]

32. Engelmann UM, et al., Predicting size-dependent heating efficiency of magnetic nanoparticles from experiment and stochastic Néel-Brown Langevin simulation. Journal of Magnetism and Magnetic Materials, 2019 471: p.450–456.

33. Zhang J, et al., Quantification and biodistribution of iron oxide nanoparticles in the primary clearance organs of mice using T1 contrast for heating. Magnetic Resonance in Medicine, 2016: p. n/a-n/a.

34. Hedayati M, et al., An optimised spectrophotometric assay for convenient and accurate quantitation of intracellular iron from iron oxide nanoparticles. International Journal of Hyperthermia, 2018, 34(4), pp.373–381. [PubMed: 28758530]

35. Ring HL, et al., Ferrozine Assay for Simple and Cheap Iron Analysis of Silica-Coated Iron Oxide Nanoparticles. ChemRxiv, 2018.
Figure 1:
Schematic of adiabatic, and non-adiabatic heating curves during the heating (coil on) and cooling (coil off) phases of a heat experiment. Common fitting time-frames to calculate SARv include: Time-Rise, Box-Lucas, Steady State, and Decay.
Figure 2:
Effect of variation in SAR$_V$ and RMSE using Box-Lucas and Time-Rise fitting as a function of SAR$_V$. Red circles indicate higher variation of SAR$_V$ and RMSE caused by measurements performed in a sub-optimal heating range. The blue square indicates the range with a consistent SAR$_V$ despite fitting method.
Figure 3:
Evaluation and start time impact on SAR\textsubscript{V} and RMSE. Plots A and B show the impact of evaluation time on SAR\textsubscript{V} and RMSE when the start time is selected at 5s after heating. The SAR\textsubscript{V} values are compared with the average SAR\textsubscript{V} across all 12 fitting parameter scenarios. A higher impact in both variation of SAR\textsubscript{V} and RMSE is observed with Time-Rise fitting. Figure 3C evaluates the percentage difference in SAR between (1) SAR\textsubscript{V} calculated excluding the first 5 s and (2) SAR\textsubscript{V} calculated including t=0–5 in the selected data range. Similarly, 3D evaluates the percentage difference in RMSE between these 2 cases. In this case, Box-Lucas fitting is observed to have larger variations in both SAR\textsubscript{V} and RMSE than variations observed with Time-Rise fitting.
**Figure 4:**
Repeatability of SAR\textsubscript{V} measurement. Plot A shows the standard deviation of the replicate measurements as a function of evaluation time (using 5s start time) for SAR\textsubscript{V} < 0.1 W/mL. Plot B right shows the standard deviation of the replicate measurement as a function of evaluation time for SAR\textsubscript{V} > 0.1 W/mL. An increase in variation is observed at lower SAR\textsubscript{V}. Furthermore, the R\textsuperscript{2} optimization results in more variable results than a statically set evaluation time.
Table 1:

Recommended Parameters to Report *

| SAR, (W/mL) | Recommendation |
|-------------|----------------|
| < 1         | For high noise, Time-Rise use RMSE to guide selection of the evaluation time, and report criteria used to select the time interval for analysis of heating data |
|             | For low noise - Either Method |
| 1 – 10      | Both methods work, report selected method and use RMSE to guide selection of evaluation time and start time |
| > 10        | Box-Lucas, use RMSE to guide start time; expect inaccuracy if dataset is shortened due to physical sample constraints (i.e. boiling). |

* SARv ranges given are specific to Bordelon et. al. & Etheridge et. al. and will vary based on heating losses in the experimental setup.
Table 2:
Recommended Parameters to Report

| Sample | Experimental Setup | Calculation Parameters | Reported Calculations |
|--------|--------------------|------------------------|-----------------------|
| Volume (≥ 2 mL) | Coil (geometry and # of turns) [14, 26, 33] | Fitting method | SAR, or SLP |
| Sample holder | Temperature probe (type & placement) [14, 16, 17] | Start point selection | RMSE of fit |
| Matrix (tissue, cells, or solution) | Insulation | Evaluation time | Standard deviation of replicate measurements |
| IONP concentration [15] | Sample placement in coil [33, 25] | Number of replicate measurements | replicate measurements |
| Method of ε₀ determination [34, 35] | Magnetic field frequency | Uniform field VOI | |