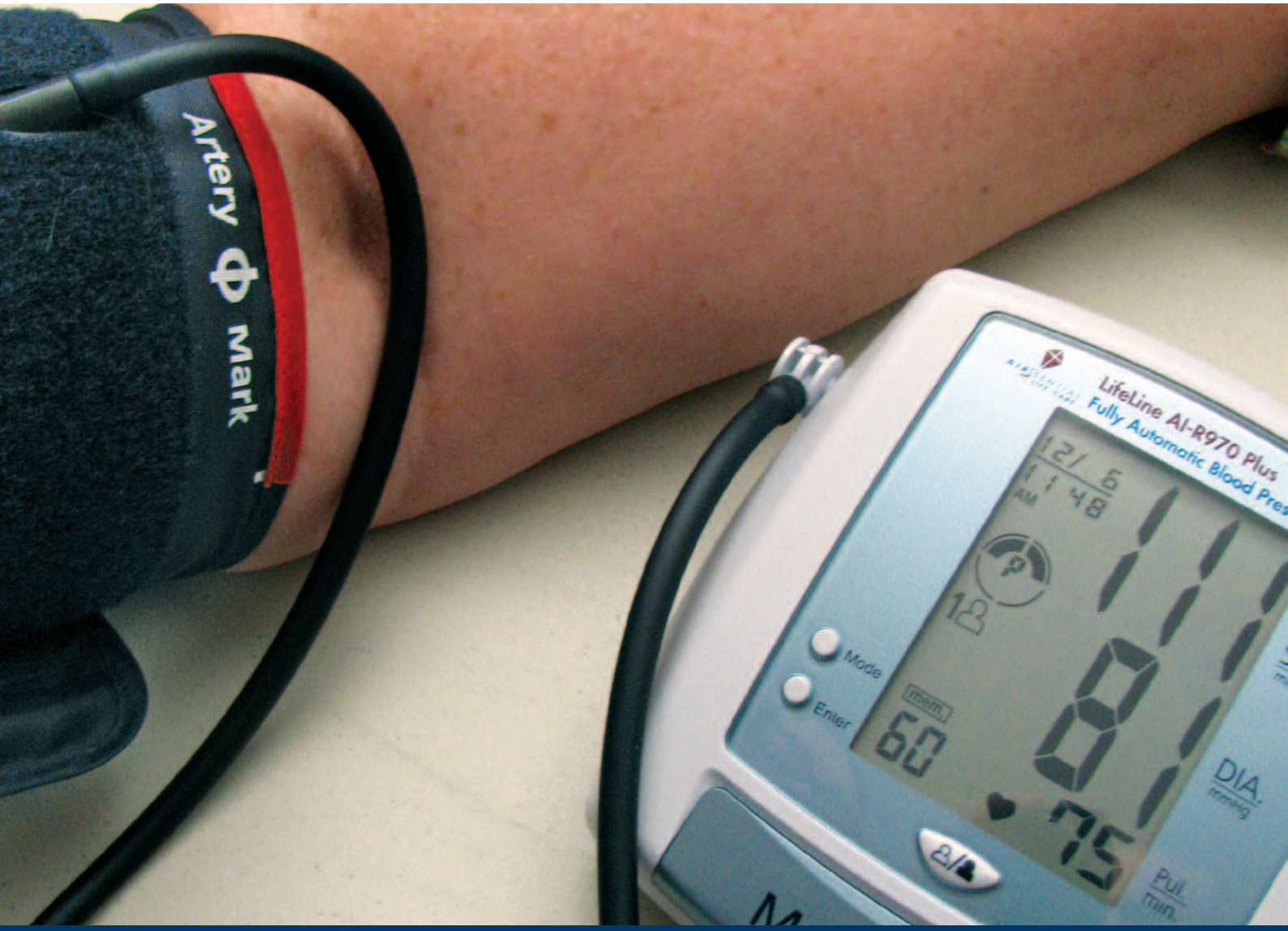
No 45 December 2010

The Australian Metrologist

The official journal of the Metrology Society of Australia



A global risk-based guideline for managing weighing systems page 11

Lack of sphygmomanometer calibration causes over- and under-detection of hypertension page 13

The Australian

Metrologist

No 45 December 2010

The official journal of the Metrology Society of Australia

The Australian Metrologist **is the journal of the Metrology Society of Australia, an Association representing the interests of metrologists of all disciplines throughout Australia.**

www.metrology.asn.au

Contents

Editorial 4

President's column 5

Letter to the editor 6

Metrology Society of Australia Award 7

Around the states 9

A global risk-based guideline for managing weighing systems 11

Lack of sphygmomanometer calibration causes over- and under-detection of hypertension 13

Interview with Dr. Martin Turner 18

Published twice a year by

CAMBRIDGE PUBLISHING
A Division of Cambridge Media

10 Walters Drive, Osborne Park WA 6017
www.cambridgemedia.com.au

Copy Editor Rachel Hoare
Graphic Designer Mark Orange

Advertising enquiries to Simon Henriques, Cambridge Publishing
Tel (08) 63145222 Fax (08) 63145299
simonh@cambridgemedia.com.au

ADVERTISING
Advertising that appears in the *The Australian Metrologist* conforms to the standards required by the Metrology Society of Australia, but endorsement is in no way implied by the publishing of said material. All advertising enquiries should be directed to the publisher, Cambridge Publishing.

EDITORIAL POLICY
All materials relevant to the practice of metrology are welcomed in the journal. Non-original material submitted must identify the source and contact details of the author and publisher. The editor reserves the right to refuse submissions. Contributors may be contacted regarding verification of material. Opinions expressed in *The Australian Metrologist* do not necessarily represent those of the Metrology Society of Australia. Material in this journal may be reproduced with prior approval of the editor.

MEMBERSHIP ENQUIRIES
Membership is available to all appropriately qualified and experienced individuals. Associate membership is also available. Contact either your state coordinator or the secretary, or visit the MSA website.

MSA Membership Fees

Fellow	$45 Annual Subscription
Member	$45 Annual Subscription
Associate	$45 Annual Subscription

Editorial

WALTER GIARDINI

A "light" issue of TAM for Christmas 2010, but some very topical issues for you to consider. Almost all laboratories have balances for weighing all sorts of things. Mass and temperature are probably the two most ubiquitous measurements we make. In this issue a short paper outlining an approach to considering what is needed and then selecting a balance which is just right for the job. Accurate enough, but not unnecessarily so.

We are reprinting a paper which MSA member Martin Turner presented at the 2005 conference in Canberra (remember!!, is it already 5 years ago?..yes). The paper presents the conclusions of some studies, and asks some questions about calibration and traceability (and hence reliability) of measurements of a person's blood pressure at the GP's. Many of us would be concerned to measure (say) temperature in our laboratories with an uncalibrated thermometer, but how about our blood pressure, probably the most commonly health/medical parameter measured in our community. Martin's paper was thought provoking in 2005, and with the paper we are also publishing an interview with Martin to bring us up to date. Is the issue still important? Have things improved in the past 5 years?

Lord Kelvin said (in 1883): "I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind". Now, over a century later, in a world in which the power and subtlety of technology in supporting human society, and life on this planet could scarcely have been imagined in Lord Kelvin's day is it still enough to say that "measurement" is enough? Or should we now demand accurate, reliable – and especially – traceable measurement. Martin's paper and interview are still relevant now, and he hints at other physical measurements (such as intraocular pressure) where rigorous metrology may have a bigger contribution to make.

As well as this there are state reports, a call for nominations for the MSA award and of course President Daniel Burke's column. We hope you enjoy this edition of TAM, and of course a Happy Christmas and great New Year in 2011 to all members, families, friends and colleagues in Metrology from the whole editorial team at TAM.

Electronic submission of manuscripts to the journal

The Australian Metrologist welcomes authors to submit all articles, whether a letter to the editor, conference report or an original article to be peer reviewed, via our web-based Manuscript Management System.

Steps to submission and publication

- Go to either the publisher's website, www.cambridgemedia.com.au, or to the TAM section of the MSA website.
- Click on Manuscript System.
- Create an account when using the system for the first time. Enter your personal and professional details, please complete all fields. These will be retained for future enquiries and submissions.
- Login.

Author guidelines are available on both the MSA and Cambridge Publishing's websites.

Submitting an article

Step 1 – Choose the Australian Metrologist, type the name of the article, choose the category of article and, if applicable, type in the abstract.

Step 2 – Add co-author details (all fields) if applicable.

Step 3 – Upload files. Please ensure your document contains the required information and is formatted according to the author guidelines. **PLEASE NOTE THAT AUTHOR DETAILS SHOULD ONLY BE ON A SEPARATE TITLE PAGE IF SUBMITTING AN ORIGINAL ARTICLE.**

Step 4 – Add any comments for the editor.

Step 5 – Review your information then click submit.

Once submitted, the manuscript is reviewed by the editor and, if applicable, sent for peer review.

Peer review

Peer reviewers will be asked to review manuscripts using the online Manuscript Management System.

President's column

DANIEL BURKE

Although the euphoria of the first issue of our magazine in the new format has levelled off to an exciting buzz, there is still enough zing remaining to start to raise some important topics for wider discussion. This magazine forms an essential pillar for communication with members and the wider community by publishing news that has a different slant from our website and eNews.

We held the Annual General Meeting on the 12th October by teleconferencing from NATA's offices in NSW, VIC, QLD and SA (see some pictures on following pages). Our Treasurer's report shows that our financial position continues to be sound with sufficient resources to continue our projects and to hold the next conference in October 2011. Randall Anderson has been our Treasurer now far slightly more than the 6 years specified in our constitution and so we need a new volunteer for that position before the next AGM in 2011. So if would like to contribute your services, please contact Randall to learn more about this important role or view the role description on our website under Members>Constitution.

Also at the AGM, I raised the topic of a new way to fulfil our stated aim of improving professional recognition for metrologists. I proposed that we consider that the MSA as a professional society may confer the title of "Chartered Metrologist" on suitably qualified and experienced members. This would then serve as a communication to the wider community that this person has special knowledge and experience in measurement that may be used for solving measurement problems or for auditing measurement practices. The legal implications and resources necessary will be investigated and reported to the members at the next AGM.

On the big screen, Sydney members discuss the agenda (President Daniel Burke at extreme right of screen)

In the conference room of the new NATA offices in Abbotsford in Melbourne, members participate in the networked meeting (thanks to the NATA national audio-video network system)

Keith Fordham, John Robertson and Randall Anderson enjoying a pre-meeting drink.

On a lighter note, it is drawing close to the festive season so keep an eye out for your regional Christmas party. I urge you to join your colleagues and have a convivial evening with your regional contact. If you are in a remote location and would have difficulty travelling to the regional Christmas party, you are welcome to organise members in your area for an MSA Christmas party that may be eligible for financial support. If you want to do this, please contact me directly at the email address on our website under Members>National Team.

Happy Christmas and a prosperous New Year to all.

Letter to the editor

Dear Editor

Rounding Decisions

In the June 2010 issue of TAM Robert Crawford asked for ideas for rounding results which are at the mid-point between successive rounded values.

The desirable property of such rounding is that it be unbiased. Because an observation is in fact the centre of a probability distribution of real values, the true value may be above or below the observed value, so is equally likely to actually be required to be rounded up as down. So rounding rules should, in the ideal case, assign the rounding direction by a random choice. A close approximation to this is that half such roundings be up and half down.

A half century or so ago when I began my metrology career we wrote readings in a book and calculated averages without any electronic assistance. We had a simple rule for rounding: we rounded down to an even number and up from an odd number. As an example, if the average was 10.5 we recorded the result as 10, if the average was 11.5 we recorded 12. In some rare circumstances this my have introduced a small bias, but in the long run it will certainly give a random assignment and was easy to do.

The method as described works only for numbers rounded to an integer, so how could it be expanded for any rounding quantity? The answer lies in the fact that the result of any rounding is an integer multiple of the rounding amount. For example if your reading is rounded to the nearest 0.05, then the rounded result divided by 0.05 is an integer number (a number without a fraction or decimals). So you could easily apply the above rule to the multiple, then use the multiple to arrive at the rounded result.

As I was writing this article I realized that this method is the same as that quoted by Robert from AS 2706, "…choose the rounding that is the product of twice the rounding interval and an integer.", but the writer of the standard had managed to make a simple rule somewhat opaque. In both cases marginal values are rounded to the even multiple of the rounding interval.

This topic brings out the issue of validation, which is checking that a calculation method or program produces correct answers. You need to determine how your process behaves to find out if there is any problem with rounding. There is a tendency for us to accept automatic systems and well-known programs as correct, but we should remember that not only is it possible for programs such as Excel to have flaws (yes, it has happened), but we may have made errors in our coding, including mistaking the exact way in which a function operates when it works as intended.

Which brings the discussion back where we started, because the problem of whether rounding is done correctly could be seen as a validation problem: does your spreadsheet or other calculation program give you appropriate rounding? If you can demonstrate that it does, no further action is required; if it does not then you need to come up with a response. Solving the problem within a spreadsheet will be straightforward, within a data acquisition system could be a bit more murky.

Are there any means to rectify the effects of biased rounding? I can think of some. First, if you are able to collect more digits than are required you could do the rounding by hand and use the method of AS2706. Second, if you can record the data collected by an automatic system you could examine the data and calculate the bias introduced into averages and correct for it. Thirdly if the problem occurs in a calculation program such as Excel you could adjust the calculation code or enter more digits.

I found that my Excel spreadsheet program rounds all midpoint values upwards, so 1.5 becomes 2 and 2.5 becomes 3. I found it easy to write a small test program in Excel which performs the "round to an even multiple" process in a few steps, and another using the pseudo-random number generator in Excel to make a random assignment of rounding direction. But Excel also rounds 1.49 down to 1 and 1.51 up to 2, so if you can obtain extra places in your observation you can get very close to perfect rounding.

How serious is the rounding problem? When thinking about how much effort to put in, you should consider what impact a problem might have relative to your measurement uncertainty. If you are rounding to integer numbers from values with one decimal place, then just one in ten of a widely distributed sample will fall on the middle value between successive integers, and if your process rounds all mid-point values in the same direction then just half of these will go the wrong way, so on average you will have an offset of one twentieth of your least significant digit, that is, 0.05. If you round from two decimal places then just one in two hundred will be rounded the wrong way.

The uncertainty due to rounding is a component with a rectangular distribution of width equal to the rounding. In my example (rounding to integers), this gives a standard uncertainty of 0.5 divided by root-three, which equals 0.29. The 95% confidence component is about twice this, so just this component alone will be bigger than any rounding bias. The worst case would be every reading falling on the same midpoint value in which case the true offset would be half a digit from the actual average, so even then the offset would be within the uncertainty. Just expanding your uncertainty a bit may be the most efficient way of handling the problem.

So my recommendations for addressing rounding when the result falls midway between two possible rounded values is as follows.

1. Look at the part of your data system where final rounding takes place. If this occurs within a program or electronic system, test to see what rounding protocol is used.
2. If the final rounding is done manually, round in conformance with AS2706, that is, round mid-interval values so that the rounded result is an even multiple of the rounding interval.
3. If rounding of a program or electronic system is found to conform with AS2706, no action is required in your calculations. But whether it does or not, document your findings so that you can justify whatever procedure you choose to use.
4. If the rounding does not conform, consider what remedial processes could be instituted.
5. Calculate the probable effect of biased rounding on your results. Compare the effect with your final uncertainties, concentrating on probable worst cases.
6. Decide whether any corrective action is possible and justified. If it is not introduce a component into your uncertainty calculations to account for it. Document this process for presentation to your assessor..

The lesson to be learned from this is that we must always keep in mind the goal we should be aiming for, which is to obtain numbers that best represent the measurement we have made. Conformance with standard specifications is not the main game and should be adhered to only where practicable and where this gives the best result. But if conventional guidelines are not followed you must be diligent in documenting what you did and why you did it.

Jeffrey Tapping
jeffrey.tapping@gmail.com

***The Australian Metrologist* welcomes contributions to the letters page, comments, opinions, questions, etc.**

Metrology Society of Australia Award

The Metrology Society of Australia Award recognizes achievement and excellence in Australian metrology and the contribution metrologists make to the Australian community. Metrology is the science of measurement. Membership of the MSA includes scientists, engineers and technicians working in government and industry from all fields of measurement in Australia and overseas.

The MSA Award is presented biennially at the MSA conference dinner. In 2011 this will take place in Victoria on Oct 19 to 21.

Nominations are now invited for this award. Only members of the Metrology Society of Australia are eligible. Members may self nominate or nominate another member using the nomination form.

The award is for work completed, or that has gained scientific or industrial recognition, in the past five years and which has contributed to the Australian economy. The work must fall into ONE OR MORE of the following categories.

BASIC RESEARCH: Original research directed towards the significant improvement of fundamental measurements, the accuracy of derived units or fundamental constants. Solutions to difficult measurement problems, work that has FUNDAMENTAL importance to the development of measurement, the application of new or existing science and mathematics to new measurement applications, including the development of new instruments, techniques or methods for reducing uncertainty.

DEVELOPMENT: The development of new instruments, measuring techniques or systems for Australian industry, including the design of prototypes, testing, characterisation and product manufacturing. For example the development of a new thermometer or an inline automatic inspection system.

APPLICATION TO INDUSTRY: The use of new or improved measurement science and technology in Australian industry to increase quality, productivity and competitiveness. For example, the use of new sensors to control production processes or the application of statistics for scheduling recalibration systems.

SELECTION PROCESS: The Award judges will be a sub-committee of the MSA National Management Committee. The judges will use criteria such as; degree of innovation, significance of the work, potential or real cost savings, stage of development, potential for application in other fields or industries; quality of supporting material and testimonial evidence supplied.

The Award judges are bound by confidentiality agreements, ensuring complete confidentiality of submitted material.

Application form overleaf

MSA AWARD APPLICATION FORM

To Nominate, please fill in the entry form below and send it to:

The Secretary,
Metrology Society of Australia
C/- National Measurement Institute
PO Box 264
Lindfield NSW 2070

Name of Nomination: .....

Address: .....

Telephone: .....

Fax:.....

Email: .....

Concise description of work on which the nomination is based:

.....

.....

.....

.....

.....

.....

Nominated by: .....

If self nominated please provide contact information below:

.....

.....

.....

.....

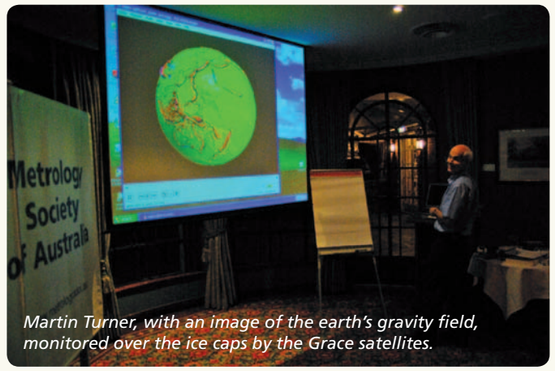
Signed: .....

Printed Name (if different from nomination):.....

Date: .....

Do you wish the submitted material to remain confidential? Y/N:.....

Around the states



Martin Turner, with an image of the earth's gravity field, monitored over the ice caps by the Grace satellites.

Peter Bradley from Zarlink Semiconductor speaking at the joint EA/IEEE/IET/MSA talk.

NSW/ACT

Since the TAM 44 report, a number of events have occurred. We had MSA representation at the NMI/CSIRO Open Day on Septenmber 19 – I thank the MSA volunteers Daniel Burke, Martin Turner, Maree Stuart, Stephen Dain, Thomas Hagen, Walter Giardini and Vasu Balakishnan who manned the MSA stand – we received 5 applications. Walter also filmed for the MSA – a great initiative. We have had a co-hosted talk at Engineers Australia on 28th October, which was a joint EA/IEEE/IET/MSA talk by Dr Peter Bradley, from Zarlink Semiconductor entitled, "A New Standard for Wireless Medical Body Area Networks IEEE 802.15.6". This talk is part of the ongoing annual reciprocal talk arrangement we have with Engineers Australia. It was a fascinating talk on a very difficult and ambitious technology, by a major player in this developing area. Next year we will find the speaker (ANY SUGGESTIONS?) and host the event. In Sydney we also had the MSA AGM at NATA in Sydney on 12th of October.

On a personal note, I will stand down as NSW Contact in the near future, handing over to Thomas Hagen of the NMI. I take this opportunity to thank all the NSW Committee members that have assisted me as the co-ordinator in NSW since August 2007 (Maree Stuart, David Hayles, Thomas Hagen, Fred Emms and

Fred Emms discussing his compiled historical temperature records at the climate Change talk in June.

former committee members Martin Turner, Veronica Vamethevan, Lee Macer-Wright, Stan Brulinski and Dunia Sukkar).

I will continue to be on the NSW committee for some time, but intend to concentrate on the other role I have had since the MSA2009 conference, that of National MSA Secretary. It has been a most interesting and rewarding time in NSW and we have some great lab tours and events and I have met many interesting people. I also thank my former NSW Chief, Daniel Burke, who I served with ont the NSW committee in 2006-2007. I welcome Thomas to the role, which I'm sure he will fill with distinction!

Paul Pokorny, Paul.Pokorny@measurement.gov.au

VIC/TAS

The MSA 2011 conference committee has been hard at work and selected the Deakin Management Centre in Victoria for the venue of the next MSA conference, on the 19th to 21st October next year. The Deakin Management Centre is located near the thriving regional city of Geelong to the South West of Melbourne, and promises to be an excellent venue for our next conference (see photos this page).

The overall theme has been chosen to be "A Climate for Measurement" so now is the time to start thinking about your papers and submissions for the conference. We are nearly at year's end for 2010, and those weeks and months will flash by, so consider what you would like to present and share with your fellow metrologist!!!

Any ideas?, questions?, suggestions?, contact any of the committee members...

neville.owen@measurement.gov.au,
keith.fordham@mt.com,
randall@auspressurelab.com.au,

Neville Owen, Neville.Owen@measurement.com.au

QLD/NT

Tuesday evening, 24 August 2010, Queensland member Ian Kendall hosted a site visit to the QR National Redbank Workshop which was attended by Queensland members.

The tour commenced with a brief safety induction and some light refreshment.

The site tour viewed three business areas; wagon manufacture, coal & freight locomotive overhaul and passenger train overhaul, as well as the various component services supporting these activities. There are approximately 900 hundred staff on site.

The wagon fabrication and assembly shops currently have two major contracts for stainless steel coal wagons for the Hunter Valley and Central Queensland coal basin.

The locomotive facility performs both deep maintenance with removal of all components including the superstructure and running repairs of locomotives in service.

The passenger business likewise performs suburban electric overhauls.

These three operations are supported by components shops manufacturing or overhauling major items including diesel engines, traction motors & alternators, bogies, brakes, wheelsets, air conditioning units and interior seating and trimmings. The foundry and machine shop also supply components to all areas of operation. The painting facility paints and decals all rollingstock units.

The tour culminated in a tour of the metrology room which supports most of the dimensional calibration needs for the business and provides a diverse and interesting range of testing projects for the staff.

The tour ran for over 2 and half hours and was enjoyed by all who attended.

Geoff Barnier, Geoff.Barnier@deedi.qld.gov.au

A global risk-based guideline for managing weighing systems
GWP® (Good Weighing Practice™)

DR. KLAUS FRITSCH
Mettler-Toledo AG, Switzerland

This article presents a universal methodology to selecting and testing weighing instruments. Considering primarily the user's weighing requirements and risks, it describes a state-of-the-art strategy to ensure reliable weighing processes em-bedded in any current quality management system.

Selecting of a Weighing Instrument – The Concept of "Minimum Sample Weight"

"I want to buy an analytical balance with a readability of 0.1 mg, because that is the accuracy I need for my application."

Statements like this are often heard when selecting a balance for a particular application. This is a misconception, for the simple reason that the readability of an instrument is not equivalent to its weighing accuracy.

There are several properties, quantified in the specifications of the weighing instrument, which limit its performance. The most important are repeatability (RP), eccentricity (also known as off-centre-loading) (EC), nonlinearity (NL) and sensitivity (SE). How do they influence the performance, and hence, the selection of a weighing instrument?

To answer this question, the term "weighing uncertainty" must first be discussed. Uncertainty is defined as a parameter which expresses the dispersion of the values of a measurement. The weighing uncertainty, i.e., the uncertainty when an object is weighed, can be estimated from the specifications of a balance (typically the case when selecting a balance for a specific application), or from test measurements with the weighing instrument (typically the case when carrying out a calibration). The essential influences can be combined according to statistical methods to obtain the weighing uncertainty. While the absolute weighing uncertainty of a balance increases with increasing sample mass, the relative weighing uncertainty $U_{rel} = U_{abs}/m$, usually expressed in %, increases with decreasing sample mass.

For electronic balances, the relative weighing uncertainty as a function of the sample mass can be generally separated into two distinctive regions with a transitional region inbetween. Region 1 with "small" sample masses (small compared with the capacity of the instrument) is indicated yellow in Figure 1. In this region, repeatability is the decisive contribution factor to weighing uncertainty, whereas the other balance parameters essentially do not contribute. Region 2 with "large" sample masses (sample masses close to the capacity) is indicated green in the figure. In this region, sensitivity (and partially also eccentricity) is dominant while repeatability generally only contributes little to the measurement uncertainty. Moreover, for a majority of laboratory balances nonlinearity hardly contributes a significant part to uncertainty, as its relative uncertainty, over the entire range of sample mass, is smaller than any other contribution.

With the knowledge of the weighing accuracy required for an application, the essential selection criterion for a weighing instrument can be formulated: The relative weighing uncertainty when weighing the smallest sample must be smaller than, or equal to, the accuracy required (A_{req}) by the user's application. This criterion defines the so-called "minimum sample weight", or simply "minimum weight", which is the smallest sample mass satisfying the required accuracy. Consequently, for a particular application with a specific accuracy requirement, the smallest sample weighed on a balance must be equal or larger than the respective "minimum weight" of this balance. As the "minimum weight" is usually a small weight (compared to the capacity), the associated weighing uncertainty is mainly governed by repeatability. Consequently, the "minimum weight" can be approximately calculated as

$$m_{min} = U_{rel}/A_{req} \approx U_{RP}/A_{req} = k \cdot s_{RP}/A_{req},$$

where s_{RP} is the standard deviation of a defined number of repeated weighings (e.g. 10 repeated weighings), and k the expansion factor (usually 2).

Routine Testing of Weighing Instruments

"Measuring equipment shall be calibrated and/or verified at specified intervals […] against measurement standards traceable to international or national measurement standards."

ISO 9001:2008, 7.6 Control of Monitoring and Measuring Devices

Most likely, the majority of all samples being weighed on laboratory weighing instruments satisfy the condition of being "small samples", i.e., samples with a net mass considerably smaller than the capacity of the weighing instrument. When discussing the relative uncertainty versus sample mass, it has already been said that weighing uncertainty is governed by repeatability, if a small sample is weighed. Consequently, with the majority of weighing processes, repeatability is the most important contribution to uncertainty. Besides the sensitivity test (which assesses the balance for any critical systematic deviation and which usually is carried out near the upper end of the weighing range where sensitivity

dominates weighing uncertainty), it is the repeatability test which is essential for confirming that the balance fulfills the accuracy requirements of the weighing process under consideration. The repeatability test should be carried out in the lower region of the balance range, e.g. with a test weight of 5% of the capacity, as it dominates weighing uncertainty here. Nonlinearity is not recommended to be tested by the user at all, as its influence on weighing uncertainty is relatively insignificant with any model of laboratory weighing instrument. In any case it is taken care of when the weighing instrument is calibrated by authorized personnel.

It is good practice to relate the frequency of user tests to the risk of the weighing application. It is assumed that the more stringent the accuracy requirements of a weighing are, the higher the probability becomes that the weighing result does not meet the accuracy requirements. In this case, the test frequency should be increased. Similarly, if the severity of the impact of an error in weighing increases, the tests should be performed more frequently. That way, a higher impact is offset by more frequent tests, thereby lowering the likelihood of occurrence of the impact, and hence, offsetting the increase of risk that otherwise would occur.

The frequencies for the test of all properties might thus extend from daily for very risky applications (however mostly automatic tests, see below), over weekly, monthly, quarterly, twice a year to yearly (e.g. calibration by authorized personnel).

Instruments with Automatic Test and Adjustment Features

Adjustment mechanisms built into weighing instruments consist of one or more reference weights, and a loading mechanism that is actuated either manually or automatically. Such a mechanism allows the convenient testing or adjusting of the sensitivity of the weighing instrument. Because the built-in weight cannot be lost, cannot be touched and is kept in a sheltered place inside the instrument, this concept has advantages over testing or adjusting with an external weight, which is vulnerable to damage, dirt and other adverse effects. It can also allows the substantial reduction of the frequency of such tests or adjustments with external reference weights.

If a weighing instrument features such an adjustment mechanism, it should be (frequently) used, as it is a procedure that requires little to no effort, with the exception of a short interruption of use to the instrument. As a consequence, routine tests of sensitivity with external reference weights may then be performed less frequently.

Conclusion

By implementing GWP® (Good Weighing Practice™) as a methodology to provide a risk-based life cycle approach for evaluation, selection and routine testing of balances, measurement errors can be reduced and reliable weighing processes can be realized.

For a specific weighing process, the so-called "minimum sample weight" of the balance for the accuracy required must be smaller than the smallest sample expected to be weighed by the user. The recommended test frequencies for user testing are increased with higher accuracy (i.e., more stringent requirements) and with increasing severity of impact. On the other hand, for less stringent process requirements and reduced risk, test efforts can be reduced accordingly. This strategy reflects current thinking about implementing a risk-based approach to instrument qualification.

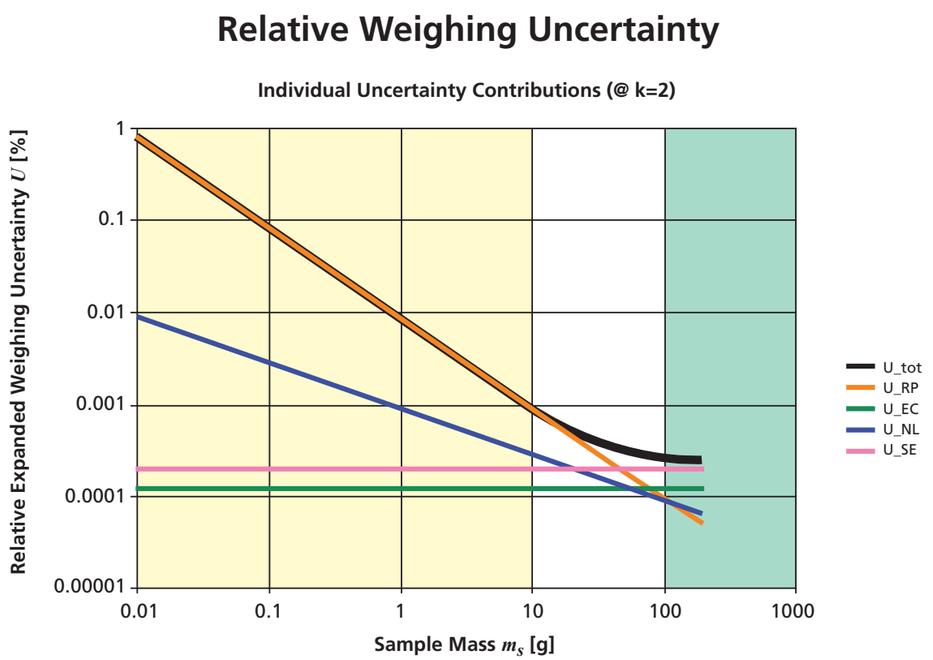


Figure: Relative weighing uncertainty versus sample mass (with zero tare load) of an analytical balance with a capacity of 200g and a readability of 0.1g (U_tot, thick black curve). The contributing components to uncertainty are also shown: repeatability (U_RP, orange), eccentricity (U_EC, green), nonlinearity (U_NL, blue) and sensitivity offset (U_SE, pink). Uncertainties are expanded with a factor of k=2. Repeatability dominates uncertainty in the yellowish region, sensitivity or eccentricity in the greenish region.

Lack of sphygmomanometer calibration causes over- and under-detection of hypertension

MARTIN J TURNER, LES IRWIG, ALEX J BUNE, A BARRY BAKER, PETER C KAM
University of Sydney, NSW 2050, Sydney, Australia

Abstract

Sphygmomanometers used in primary healthcare are often not adequately maintained and traceably calibrated. The effects of systematic and random measurement error on the detection of hypertension have not been well quantified.

We performed Monte Carlo simulations to estimate the effects of sphygmomanometer calibration error and random inter-day intra-individual blood pressure (BP) variation on the detection of hypertension in Australian adults 18 years and older. The distribution of uncalibrated sphygmomanometer errors was obtained from a recent UK survey of 1462 mercury and aneroid sphygmomanometers. Inter-day intra-individual BP variation was estimated from the Canadian Heart Health Survey (1998–1992). The distribution of BP in Australian adults was estimated from the Australian National Nutrition Survey (1995-1996). Primary care clinicians were assumed to measure BP twice during each visit, and the number of simulated visits was varied from 1–16.

After three visits (a total of six BP measurements) approximately 70,000 Australian adults with systolic hypertension (SBP > 140 mm Hg) and 60,000 with diastolic hypertension (DBP > 90 mm Hg) are undetected, and approximately 84,000 and 150,000 adults respectively may be incorrectly classified as hypertensive due to sphygmomanometer error alone. Of all 18–24 year old females thought to have diastolic hypertension after three visits to doctors using uncalibrated sphygmomanometers, only 25% were truly hypertensive. Sphygmomanometer calibration increases this rate to 42% after three visits.

If uncalibrated sphygmomanometers have errors similar to those in the UK, they are a preventable cause of substantial misdiagnosis of hypertension in Australia, which is not corrected by averaging several measurements. Further studies are required to quantify the performance of Australian sphygmomanometers and to estimate the cost-effectiveness of calibrating sphygmomanometers traceably.

Keywords: Sphygmomanometer, blood pressure measurement, random variability, systematic error, hypertension, calibration, quality control.

Introduction

Sphygmomanometers used in primary healthcare are often not adequately maintained and calibrated, and yield erroneous measurements[1-11]. Aneroid and electronic sphygmomanometers, which are likely to replace mercury manometers[12], may be less accurate when uncalibrated than uncalibrated mercury manometers[1-3,5,7,9-11]. Uncalibrated sphygmomanometers may cause systematic errors in blood pressure (BP) measurements that are not reduced by averaging several measurements.

Random BP variability comprises intra-individual variability and random measurement error and can be reduced by averaging. Rosner and Polk[13,14] suggest that two measurements per visit provide a substantial reduction in the effects of random BP variability compared with one measurement per visit, but the additional benefit of three or more measurements per visit is small. Current hypertension guidelines recommend that treatment decisions should be based on an average of multiple BP measurements made on each of several occasions[15-17].

Current guidelines for cardiovascular risk management include BP as an important factor and recommend treatment for patients whose BP exceeds recommended thresholds. For example, antihypertensive treatment is recommended for patients at moderate or medium risk and whose BP exceeds 140/90 mm Hg[16,18].

A recent analysis of UK and Canadian BP surveys indicates that the detection of hypertension is very sensitive to systematic errors in BP measurements[19]. Rouse and Marshall[8] modelled the effects of systematic sphygmomanometer errors on the positive predictive value of BP measurements for detecting hypertension in young adults, but omitted random BP variability from their model. Marshall[20] modelled the effects of random variability on the positive predictive value of BP measurements, but did not consider systematic errors.

The combined effects of inadequate sphygmomanometer calibration and random BP variability on the detection of hypertension have not been studied. We used simple Monte Carlo simulation to estimate the combined effects of systematic sphygmomanometer error and inter-day random variability of BP measurements on the over- and under-detection of hypertension in adults 18 years and older.

Methods

The cumulative distribution of the errors of 1462 mercury and aneroid sphygmomanometers reported by Rouse & Marshall[8] was truncated at ±30 mm Hg and normalised to 100%. Linear interpolation was used between the published data points. Two sets of N_{gp} random numbers were generated, where N_{gp} is the number of general practitioners in Australia[21]. Firstly, a set of N_{gp} uniformly distributed random numbers between zero and 100 was

generated and used to select an uncalibrated sphygmomanometer measurement error, e_{su}, from the cumulative distribution of sphygmomanometer errors by a numerical inverse transform method[22]. Secondly, a set of N_{gp} normally distributed random numbers with zero mean and standard deviation 0.5 mm Hg was generated to simulate the sphygmomanometer errors, e_{sc}, after calibration. A standard deviation of 0.5 mm Hg corresponds to an expanded uncertainty of ±1 mm Hg. This calibration uncertainty is readily available from accredited pressure calibration laboratories in Australia and many other countries. Each sphygmomanometer, assumed to be used by a single GP, was assigned a single uncalibrated error (e_{su}) and a single calibrated error (e_{sc}).

Systolic and diastolic blood pressures of adults aged 18 years and older were obtained from the Australian National Nutrition Survey[23]. In this survey the BP of each subject was measured twice during a single visit by a professional nutritionist trained in BP measurement. We calculated cumulative distributions of diastolic and systolic BP at pressure increments of 1 mm Hg for Australian male and female populations in seven age groups (18–24, 25–34, 35–44, 45–54, 55–64, 65–74 and 75 years and older) and normalised the distributions to 100%. The BP distributions include a component associated with intra-individual BP variability which is substantially smaller than inter-individual variability[13]. Removing the intra-individual variability component from the distributions by deconvolution[24] did not change the simulation results materially.

The following procedure was followed for each of the fourteen age and sex groups. The number of adults in Australia[25] (N_p) was divided by the number of general practitioners in Australia (N_{gp}) to estimate the average number of adults in the group (N_a) seen by each general practitioner: $N_a = N_p/N_{gp}$. A set of N_p simulated true diastolic BP values (BP_{true}) was calculated from the cumulative BP distributions by inversely transforming N_p uniformly distributed random numbers[22], and N_a true BP values were associated with each sphygmomanometer.

We assumed that a primary care clinician uses a single sphygmomanometer for all measurements and bases treatment decisions on the mean of $2N_v$ BP measurements (two measurements made on each of N_v visits). Sphygmomanometer error was treated as a systematic error common to all measurements made by that sphygmomanometer. Two simulated BP measurements, BP_1 and BP_2 were calculated for each adult:

$$BP_1 = BP_{true} + e_{sc} + e_r$$

$$BP_2 = BP_{true} + e_{su} + e_r$$

where BP_1 represents the average of N_v pairs of BP measurements made using a properly maintained and calibrated sphygmomanometer, BP_2 represents the average of N_v pairs of measurements made with an uncalibrated sphygmomanometer and e_r is a normally distributed zero-mean random component due to inter-day intra-individual variability of BP measurements. The standard deviation of e_r is $s_r = s_b\sqrt{N_v}$. Estimates of s_b for each age and sex group are described in the Appendix. The number of visits, N_v, was assigned values 1, 2, 3, 4, 5, 8 and 16.

Each simulated true and measured systolic and diastolic BP was compared with a threshold (140/90 mm Hg) and classified as normotensive or hypertensive. Thresholds of 140/90 mm Hg are recommended for initiation of treatment of patients with moderate[16] or medium[18] risk of cardiovascular disease, and for classifying isolated systolic and diastolic hypertension[17]. True and false positive and negative hypertensive values were identified. We calculated the sensitivity, the false positive rate (1 – specificity) of BP measurements for detecting hypertension, and the number of excess false negatives and positives due to sphygmomanometer calibration error.

Reported results are the means of ten simulation runs. The coefficients of variation are small and are not reported for simplicity. Systolic and diastolic results were analysed separately. Data analysis was performed using software written in Matlab (Mathworks Nattick, USA). On each simulation run the seeds of the pseudo-random number generators were reset to new values obtained from the real-time clock, so each simulation run was independent of all other runs.

Results

In 2004 there were approximately 17,000 general practitioners in Australia, 7.54 million adult males and 7.81 million adult females. The NNS survey comprised 10695 adults. Each age and sex group contained between 259 and 1155 subjects.

Eighty-four percent of the 1462 sphygmomanometers surveyed by Rouse and Marshall[8] displayed errors less than ±3 mm Hg, 90% of errors were less than ±5 mm Hg and 96% were less than ±10 mm Hg (Figure 1).

Under-detection of hypertension

Increasing the number of visits to the general practitioner increases the sensitivity (figure 2) but the incremental benefit per visit diminishes with increasing number of visits. The separation between the calibrated and uncalibrated sensitivity curves (figure 2) represents hypertensive individuals who are not detected due to uncalibrated sphygmomanometer error. When the number of visits is small the effects of uncalibrated sphygmomanometer error on sensitivity are partially masked by random variability, but the number of hypertensive individuals not detected due to uncalibrated sphygmomanometer error increases with increasing number of visits.

After three visits approximately 39,000 and 30,000 Australian males and females respectively with systolic hypertension are missed due to sphygmomanometer calibration error (Figure 3). Approximately 38,000 and 19,000 Australian males and females respectively with diastolic hypertension are missed due to sphygmomanometer calibration error (Figure 3). The lowest sensitivity is seen in young adults (Figure 2), but sphygmomanometer error causes the largest number of false negatives in middle aged males (Figure 3). After three visits sphygmomanometer error causes 27% of all false negatives in 35–44 year old females with systolic hypertension and 37% of all false negatives in 18-24 year old females with diastolic hypertension.

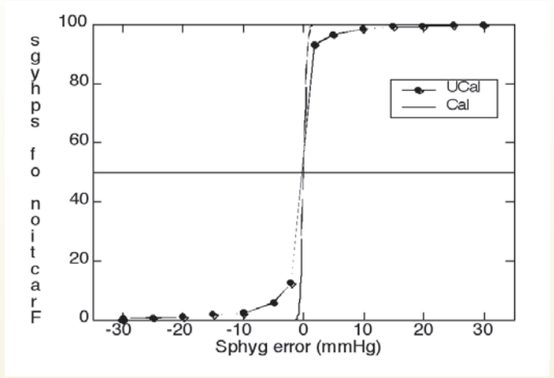


Figure 1. Cumulative distribution of calibrated and uncalibrated sphygmomanometer errors UCal: uncalibrated sphygmomanometers[8]. Cal: calibrated sphygmomanometer error (expanded calibration uncertainty: ±1 mm Hg).

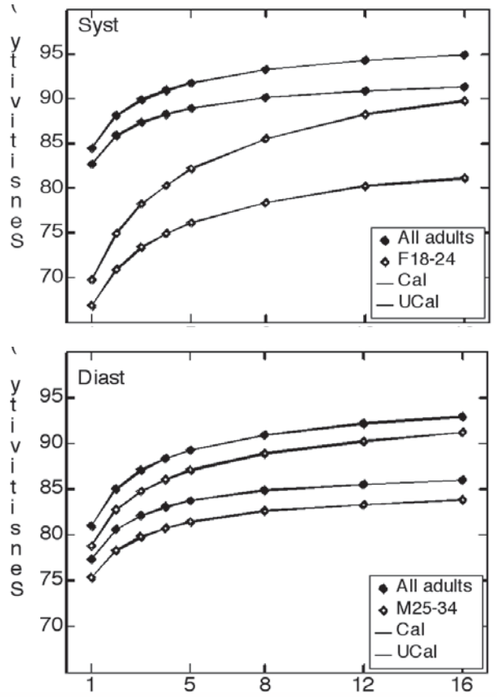


Figure 2. The sensitivity of simulated calibrated and uncalibrated BP measurements for detecting hypertension after 1–16 visits to the general practitioner. Groups with lowest uncalibrated sensitivities are indicated by open circles. F18-24: 18-24 year old females. M25-34: 25-34 year old males. Cal: calibrated sphygmomanometers. UCal: uncalibrated sphygmomanometers.

Over-detection of hypertension

Increasing the number of visits to the general practitioner decreases the false positive rate (figure 4) but the incremental benefit per visit diminishes with increasing number of visits. The separation between the calibrated and uncalibrated false positive rate curves (figure 4) represents normotensive individuals who are falsely detected due to uncalibrated sphygmomanometer error. When the number of visits is small the effects of uncalibrated sphygmomanometer error on the false positive rates are partially masked by random variability, but the number of normotensive individuals classified as hypertensive due to uncalibrated sphygmomanometer error increases with increasing number of visits.

After three visits systolic hypertension is incorrectly detected in approximately 50,000 and 34,000 normotensive Australian males and females respectively due to sphygmomanometer calibration error (fig 5). Diastolic hypertension is incorrectly detected in approximately 82,000 and 69,000 normotensive Australian males and females respectively due to sphygmomanometer calibration error (Figure 5). The highest false positive rates are seen in middle aged and older adults (fig 4), but sphygmomanometer error causes the largest number of false positives in young and middle aged males (Figure 5). After three visits sphygmomanometer error causes 63% and 50% of all false detection of systolic and diastolic hypertension detection respectively in 18-24 year old females.

Positive predictive value of BP measurements for detecting diastolic hypertension in 18-24 year old females is 25% for uncalibrated and 42% for calibrated sphygmomanometers respectively after three visits.

Discussion

This study suggests that approximately 70,000 Australian adults with systolic hypertension and 60,000 with diastolic hypertension are not detected after three visits to the general practitioner as a result of uncalibrated sphygmomanometer error. Sphygmomanometer error results in 84,000 and 150,000 Australian adults respectively being falsely detected as having systolic and diastolic hypertension respectively. Patients whose hypertension is not detected are at increased risk of heart disease, stroke and other cardiovascular diseases. Patients in whom hypertension is incorrectly detected may suffer adverse effects of unnecessary treatment.

All hypertension guidelines recommend that clinicians average multiple BP measurements to reduce the effects of random variability[16-18]. Lack of sphygmomanometer calibration, however, is not well recognised as a potential cause of systematic error in BP measurements. As the number of BP measurements increases, the negative effects of random variability decrease but the positive effects of calibration (or the negative effects of lack of calibration) become more pronounced. Systematic sphygmomanometer errors impose upper bounds on sensitivity and lower bounds on false positive curves beyond which averaging is ineffective at reducing errors in the detection of hypertension.

In Australia and the United Kingdom all medical pathology and biochemistry laboratories are formally accredited to international standards by national accreditation bodies. In the USA pathology laboratories are accredited by the College of American Pathologists. In contrast, no BP measurement guidelines recommend that sphygmomanometers should be calibrated by accredited organisations or that sphygmomanometer calibration be traceable to national or international standards. While there is rigorous quality control over laboratory measurements of cardiovascular risk factors such as lipids, there is no formal quality control over sphygmomanometers used to detect hypertension.

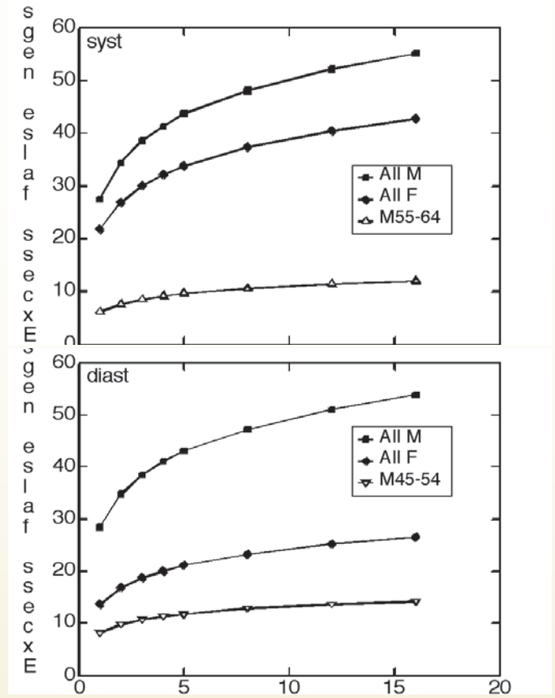


Figure 3. Hypertensive patients missed due to sphygmomanometer error. M 55-64: 55–64 year old males. M45–54: 45–54 year old males

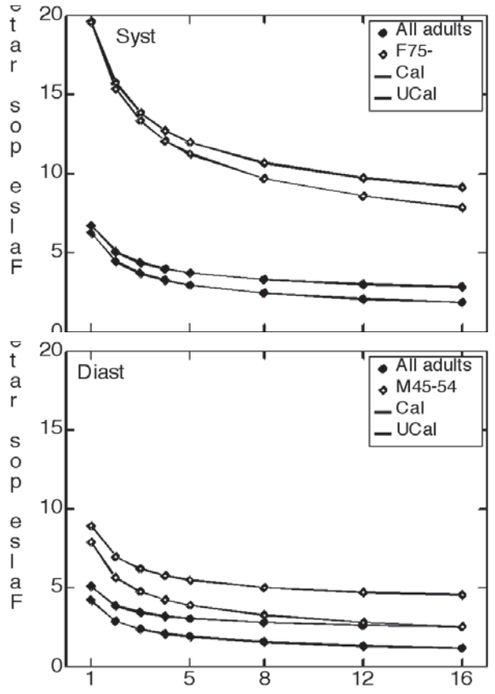


Figure 4. False positive rate (1 – specificity)

The present study suggests that the requirements of the international protocol for assessing automated sphygmomanometers[26] and the European, Australian and USA standards for manual sphygmomanometers[27,28] may be too relaxed. In the present study 84% of uncalibrated sphygmomanometer errors were less than ±3 mm Hg, 90% were less than ±5 mm Hg and 97.4% were less than ±15 mm Hg (figure 1). The protocol for assessing new blood pressure measuring devices allows a substantial number of comparisons between the reference sphygmomanometer and the sphygmomanometer under test to differ by up to 15 mm Hg. Systematic errors of up to ±5 mm Hg are permissible. Current standards for new manual sphygmomanometers[27,28] allow systematic pressure measurement errors up to ±3 mm Hg. There are no protocols or standards governing the maintenance and calibration of sphygmomanometers in use. Our simulation highlights the need for more comprehensive assessments of sphygmomanometer calibration error to improve our understanding of this source of error in the detection and management of hypertension.

Lack of sphygmomanometer calibration is a substantial and preventable cause of under- and over-detection of hypertension in adults. Sphygmomanometers should be properly maintained and regularly calibrated. Guidelines for BP measurement and hypertension management should strongly emphasise the importance of rigorous quality control of sphygmomanometers. Formal quality control procedures that include regular maintenance and calibration of sphygmomanometers by accredited organisations or individuals should be introduced.

Acknowledgements

The authors acknowledge the contribution of Dr Johan van Schalkwyk who kindly reviewed the manuscript.

The authors are pleased to acknowledge funding from The Douglas Joseph Fellowship, The University of Sydney, The Jobson Foundation, The Australian National Health & Medical Research Council (Project grants 211205 and 211001) and Dräger Australia Pty Ltd.

Appendix. Intra-individual variance of BP measurements

In the Canadian Heart Health Survey[29] trained nurses using Hg sphygmomanometers made two BP measurements in the subjects' homes and a further two measurements during a clinic visit on a different day up to two weeks later. We estimated the means and 95% confidence intervals[30] of inter-day intra-individual variances of systolic and diastolic BP from the mean of the two home measurements and the mean of the two clinic BP measurements made in this survey. Intra-individual variances were weighted using the weighting factors published with the survey data so the standard deviations are representative of the population.

The standard deviation of systolic BP grouped by age and sex correlates significantly with mean group systolic BP (R^2 = 0.918 for males, R^2 = 0.912 for females, P < 0.01). The 95% confidence intervals of the intercept of the male and female systolic regression lines do not include zero, indicating that while the standard deviation of systolic BP does depend on mean SBP, it is not directly proportional to the mean BP in the Canadian Heart Health Survey. We therefore cannot conclude that the coefficient of variation of systolic BP (standard deviation divided by mean) is independent of age and sex. Standard deviation of group diastolic BP does not correlate significantly with either mean diastolic BP of the groups or with the mean age of the groups. We therefore pooled all age and sex groups and calculated a single standard deviation for inter-day intra-individual diastolic BP.

Standard deviations used in the Monte Carlo simulations

Weighted average diastolic standard deviation was 4.99 mm Hg independent of age & sex.

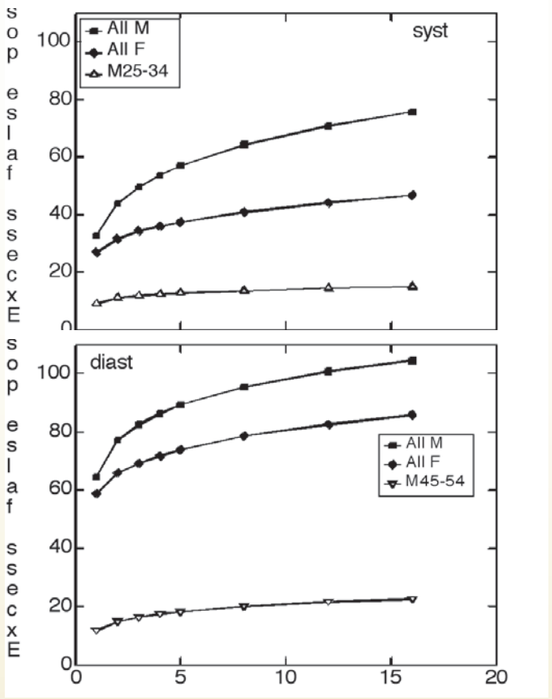


Figure 5. False positives due to sphygmomanometer error.

Weighted std dev of systolic BP:

males: $s_b = 0.1214\ SBP_{mean} - 7.8769$ mm Hg

females: $s_b = 0.1097\ SBP_{mean} - 5.9533$ mm Hg

where SBP_{mean} is the weighted average systolic BP of the age and sex group.

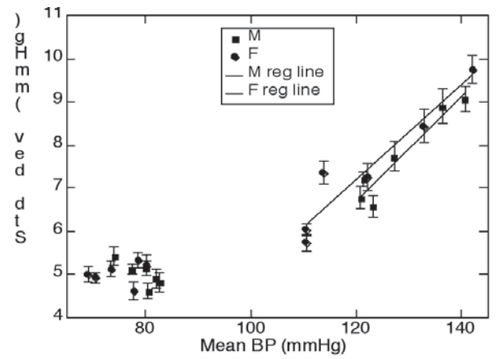


Figure 6. Standard deviations of blood pressures of each group. Solid and broken lines are regression lines fitted to male and female systolic standard deviations respectively. M: males. F: females. Error bars indicate the 95% confidence intervals.

References

1. O'Brien E, Mee F, Atkins N, O'Malley K. Inaccuracy of seven popular sphygmomanometers for home measurement of blood pressure. *J Hypertens*. 1990;8:621-34.
2. Ali S, Rouse A. Practice audits: reliability of sphygmomanometers and blood pressure recording bias.*J Hum Hypertens*.2002;16:359-61.
3. Knight T, Leech F, Jones A, Walker L, Wickramasinghe R, Angris S, Rolfe P. Sphygmomanometers in use in general practice: an overlooked aspect of quality in patient care. *J Hum Hypertens*. 2001;15:681-4.
4. BaileyRH, KnausVL, BauerJH. Aneroidsphygmomanometers. An assessment of accuracy at a university hospital and clinics. *Arch Intern Med*. 1991;151:1409-12.
5. Mion D, Pierin AM. How accurate are sphygmomanometers? *J Hum Hypertens*. 1998;12:245-8.
6. Anlauf M, Tholl U. Blutdruck-Messautomaten. Selbst mit Prufsiegel oft ungenau! [Automatic blood pressure monitors. Even with seal of approval they are often inaccurate!]. *MMW Fortschr Med*. 2000;142:34-6.
7. Ashworth M, Gordon K, Baker G, Deshmukh A. Sphygmomanometer calibration: a survey of one inner-city primary care group. *J Hum Hypertens*. 2001;15:259-62.
8. Rouse A, Marshall T. The extent and implications of sphygmomanometer calibration error in primary care. *J Hum Hypertens*. 2001;15:587-91.
9. Waugh JJ, Gupta M, Rushbrook J, Halligan A, Shennan AH. Hidden errors of aneroid sphygmomanometers. *Blood Press Monit*. 2002;7:309-12.
10. McCartney P, Crawford D. Inaccurate, leaky sphygmomanometers are still common. *Br J Gen Pract*. 2003;53:61-2.
11. Burke MJ, Towers HM, O'Malley K, Fitzgerald DJ, O'Brien ET. Sphygmomanometers in hospital and family practice: problems and recommendations. *Br Med J (Clin Res Ed)*. 1982;285:469-71.
12. O'Brien E. Demise of the mercury sphygmomanometer and the dawning of a new era in blood pressure measurement. *Blood Press Monit*. 2003;8:19-21.
13. Rosner B, Polk BF. The implications of blood pressure variability for clinical and screening purposes. *J Chronic Dis*. 1979;32:451-61.
14. Rosner B, Polk BF. The instability of blood pressure variability over time. *J Chronic Dis*. 1981;34:135-9.
15. O'Brien E, Asmar R, Beilin L, Imai Y, Mallion JM, Mancia G, Mengden T, Myers M, Padfield P, Palatini P, Parati G, Pickering T, Redon J, Staessen J, Stergiou G, Verdecchia P. European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. *J Hypertens*. 2003;21:821-48.
16. Wing L, Hankey G, Puddey I, Stowasser M, Vial J, Dart A, Duggan K, Nelson M, Boyden A, Smith J. Hypertension Management Guide for Doctors 2004. Canberra, ACT: 2003 National Heart Foundation of Australia.; 2003.
17. Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves JW, Hill MN, Jones DH, Kurtz T, Sheps SG, Roccella EJ. Recommendations for blood pressure measurement in humans: an AHA scientific statement from the council on high blood pressure research professional and public education subcommittee. *J Clin Hypertens (Greenwich)*. 2005;7:102-9.
18. Whitworth JA. 2003 World Health Organization (WHO)/ International Society of Hypertension (ISH) statement on management of hypertension. *J Hypertens*. 2003;21:1983-92.
19. Turner MJ, Baker AB, Kam PC. Effects of systematic errors in blood pressure measurements on the diagnosis of hypertension. *Blood Press Monit*. 2004;9:249-53.
20. Marshall T. When measurements are misleading: modelling the effects of blood pressure misclassification in the English population. *BMJ*. 2004;328:933.
21. Britt H, Miller GC, Knox S, Charles J, Valenti L, Henderson J, Pan Y, Bayram C, Harrison C. General practice activity in Australia 2002-03. *AIHW Cat. No. GEP 14. Canberra: Australian Institute of Health and Welfare (General Practice Series No. 14)*. Canberra; 2003.
22. Rubenstein RY. *Simulation and the Monte Carlo Method*. New York: Wiley; 1981 p38-9.
23. McLennan W. National nutrition survey, Confidentialised Unit Record File. *Australian Bureau of Statistics*. 1995.
24. Cordy CB, Thomas R. Deconvolution of a distribution function. *J Am Stat Assoc*. 1997;92:1459-1465.
25. Trewin D. Population Projections Australia 2002 to 2101. *Australian Bureau of Statistics*. Canberra; 2003.
26. O'Brien E, Pickering T, Asmar R, Myers M, Parati G, Staessen J, Mengden T, Imai Y, Waeber B, Palatini P, Gerin W. Working Group on Blood Pressure Monitoring of the European Society of Hypertension International Protocol for validation of blood pressure measuring devices in adults. *Blood Press Monit*. 2002;7:3-17.
27. AS EN 1060.1-2002. Non-invasive sphygmomanometers - General requirements. Standards Australia, Sydney. 2002.
28. ANSI/AAMI SP9:1994 Non-automated sphygmomanometers. Association for the Advancement of Medical Instrumentation. Arlington VA. 1994.
29. MacLean DR, Petrasovits A, Nargundkar M, Connelly PW, MacLeod E, Edwards A, Hessel P. Canadian heart health surveys: a profile of cardiovascular risk. Survey methods and data analysis. Canadian Heart Health Surveys Research Group. *CMAJ*. 1992;146:1969-74.
30. Dixon WJ, Massey FJ. *Introduction to statistical analysis 4th Ed*. New York: McGraw Hill; 1983 p170.

Interview with Dr. Martin Turner

Interviewed by W. Giardini

Interview with Dr. Martin Turner on his efforts to improve the situation regarding the calibration and hence the reliable accuracy of sphygmomanometers.

Q: **The paper you presented at the MSA 2005 conference in Canberra, suggests that there is a significant proportion of incorrectly diagnosed condition of high blood pressure in the community. What are the practical consequences of this?**

Moderately increased blood pressure usually causes no symptoms, but it increases the risk of cardiovascular disease (heart disease, kidney disease, stroke, arterial disease). A common reason for measuring blood pressure (BP) is to estimate the risk of cardiovascular disease (CVD). Risk of CVD is calculated from BP, cholesterol, age, sex, smoking habit, presence of diabetes and presence of changes in the ECG that indicate an enlarged heart. Lifestyle changes and treatment are aimed at reducing the risk. If BP is high then damage can also occur to some organs, e.g. large arteries, blood vessels in the eyes, brain, heart and kidneys. If BP is underestimated and hence under-treated, risk of heart attack, stroke and organ damage is increased. If BP is over-estimated and hence over-treated, the patients get drugs that they would not otherwise get and hence they suffer from the side effects of the drugs and interactions that the drugs might have with other drugs that the patient might be taking.

Q: **You say in the paper that of all 18-24 year old females who are thought to have diastolic hypertension due to sphygmomanometer error, only 25% are truly hypertensive. Do I understand correctly that you are saying that only a quarter of women who are diagnosed with hypertension actually have it?**

Yes. Very few young women have high blood pressure. When the BP of a large number of young women is measured with instruments that have the broad distribution of systematic errors that we used in this study, it is more likely that a high reading will result from a systematic measurement error than from a truly high BP.

Q: **You make a point of separating the statistics for males and females. Is the difference between the sexes in the instance of hypertension significant? Do we know why those differences exist, or can we hazard guesses?**

I don't understand all the physiological mechanisms. Women tend to have lower BP than men until menopause and their CVD risk is different (usually lower at a given age). After menopause their BP seems to rise faster than men's BP, so the prevalence of hypertension in older women is higher than the prevalence in older men. There is fairly strong evidence that the rise of BP with age is mainly due to our very high salt intake, and that women's BP is more sensitive to salt after menopause. Not everyone's BP rises with age. There is lots of inter-individual biological variability. Some people's BP (both men & women) stays low all their life no matter what they eat.

Q: **You presented this article at the MSA conference in Canberra in 2005. How serious do you think this problem of misdiagnosed hypertension is in the overall health management of Australians?**

The distribution of calibration errors we used in our study is from a UK study. We don't know what the distribution is in Australia, but there is no reason to believe it is any better. If it is roughly the same as the UK study then there are many thousands of Australians taking antihypertensive drugs who otherwise wouldn't and many thousands who should be taking antihypertensive drugs who aren't. This means loss of life and quality of life for people whose BP is under-treated and loss of quality of life and possibly loss of some lives (from adverse reactions to the drugs and from interactions between drugs) for people who are over-treated. I think in my early (unpublished) calculations I also found that there was a net increased dollar cost (in the $millions) to the PBS caused by the sphygmomanometer errors we used in our study, but these results need to be confirmed.

Q: **What part of the medical and health industry would be concerned with understanding the extent of this problem, its significance, and possible ways to improve the situation?**

GPs and physicians who diagnose, treat and monitor hypertension and CVD and risk of CVD should be concerned. Medicare and the PBS should be concerned that the right drugs get to the right people.

In 2007 Noel Bignell, Catherine Speechly (of the RACGP) and I published a paper in the Australian Family Physician entitled *"Sphygmomanometer calibration Why, how and how often?"*. We had a handful of enquiries from clinicians enquiring how to get their sphygmomanometers calibrated. I heard that one or two clinicians could not find a NATA accredited calibration lab to calibrate their Hg sphygmomanometers because the labs could not deal with Hg for OH&S reasons.

Q: **What are the barriers to sound metrology practices being understood and implemented in non-metrology fields such as this one, and what do you think are the mechanisms to address those barriers?**

Clinicians have (or perceive that they have) very little time to deal with this kind of problem. I think they prefer to deal with the patients and their medical problems and don't want to worry about measurement details. The human body is a very complicated organism, so that is probably quite reasonable. There is a lot of variability in physiological measurements, caused in part by biological variability. For example, the standard deviation of systolic blood pressure measured at intervals of few days or more is in the region of 12 mm Hg. Many clinicians think that in the face of high biological variability, some measurement error doesn't matter. The problem with that thinking is related to the difference between random variation and systematic error. Perhaps some education could help.

Many clinicians (including some of the experts who write the guidelines on BP measurement) seem to think that Hg manometers never need calibration. Metrologists know that all measurement systems need regular calibration.

The TGA tightly regulates the manufacture, importation and distribution of medical devices in Australia. Manufacturers are required to report adverse events involving medical devices to the TGA, but erroneous measurements, if detected, do not appear to constitute adverse events. The TGA appears to have no influence on maintenance and quality control of medical devices with a measurement function after they are sold.

Calibration is not the only regular service that sphygmomanometers need. They should be regularly checked for leaks and the dynamic response should be checked by the "rapid exhaust test". I think it is pretty clear that most clinicians are unlikely to willingly make the effort to get their instruments properly calibrated and serviced. That is going to have to be mandated somehow, perhaps as part of new accreditation requirements that may be coming. The Australian Commission on Safety and Quality in Health Care is considering across the board accreditation for healthcare workers.

The metrology infrastructure needs a little development to provide a good service that is easy and cost-effective for clinicians. For example, it is probably best if sphygmomanometers (and other instruments) are calibrated on site in clinicians' offices. That will get around the OH&S problem with Hg and reduce the down-time of the instruments. That means training metrologists to do the job. It needs a system similar to the licensing system that verifies scales in supermarkets.

Q: **What progress would you say (if any) has been made in improving the metrological control of sphygmomanometers?**

In Australia very little (as far as I know). In fact in the last decade or so we might have gone backwards. An uncalibrated Hg sphygmomanometer is probably better than an uncalibrated aneroid sphygmomanometer, but Hg sphygmomanometers are slowly disappearing because of the toxicity of Hg, and being replaced with aneroid and automatic electronic devices.

Aneroid manometers are known to lose calibration very quickly.

In general, automatic sphygmomanometers are not particularly good due to the methods used to detect systolic and diastolic pressures, and the lack of quality control during in-vivo validations of these devices. There is some evidence, however, that the pressure measurement subsystems in automatic sphygmomanometers are more reliable than aneroid and Hg manometers.

Q: **What is the current state of play, and are you still lobbying to improve the situation?**

As far as I can see the current state of play is not different to what it was in 2005. There is an unknown number of unmaintained, uncalibrated sphygmomanometers out there causing an unknown number of diagnostic and treatment errors.

The same situation applies to other physical medical measurements such as intraocular pressure and respiratory measurements.

I am trying to bring together the NMI and the Royal Australian Colleges of physicians and GPs to discuss the problem and possible solutions. The Chief Metrologist, Laurie Besley is interested but so far I am having difficulty getting the clinical people involved.

An article I wrote suggesting that automatic sphygmomanometer validation should be done by laboratories accredited to ISO 17025 has recently been published in J Hypertension. It was accompanied by an editorial by world experts in BP measurement, and may stimulate thought and discussion among clinical people.

Q: **Why does society put so much effort into quality control of measurements for trade, industry and science, but seems to neglect quality control of many medical measurements?**

I don't know, but it is a situation that needs to change. We have a world-class metrology infrastructure in Australia but many medical measurements, particularly physical measurements, appear to ignore that infrastructure. The consequences are many measurements with clinically significant systematic errors and high inter-device and inter-laboratory variability. That means that a person who has, for example, respiratory measurements made at one lab, can't have follow-up measurements made at a different lab. Evidence in the literature suggests that follow up measurements must be made on the same instrument in the same lab. At least one large, multi-centred human study had to run a large, separate project to minimise and control inter-laboratory differences in respiratory measurements. Good quality traceable calibration and proper quality control should make this kind of extra effort unnecessary.

A short video clip of edited highlights of this interview is being prepared for the MSA website, so look for it in January.
www.metrology.asn.au