2013 IEEE International Conference on Healthcare Informatics

Using nomograms to reduce harm from clinical calculations

Harold Thimbleby
College of Science
University of Swansea, Wales
Email: harold@thimbleby.net

David Williams
Department of Anaesthetics
Abertawe Bro Morgannwg University (ABMU) Health Board, Wales
Email: davidjwilliams@doctors.org.uk

***Abstract*—Handheld calculators and computer spreadsheets are ubiquitous and taken for granted. However in hospitals, errors in routine calculations frequently occur — for instance after making an unnoticed typing slip, such as omitting a decimal point in a drug dose calculation — and can result in patient harm. This paper is concerned with dependable calculation, and examines user tasks and technologies for safer calculations in the clinical environment. We demonstrate significant differences in complexity, speed and accuracy between alternative methods of performing calculations.**

The recent raised awareness of latent coding errors in applications designed to perfom medical dosage calculations has resulted in the introduction of national regulations that require all medical apps to meet similar standards of safety and reliability as other items of medical equipment. This paper provides evidence that general purpose calculators are also unnecessarily hazardous in the clinical environment, and should be subject to similar regulation. This paper contributes to the current debate about the use of computer systems to improve healthcare, and argues that "the latest IT" does not automatically confer benefit: its effectiveness should be empirically evaluated like any other medical intervention.

The combination of simple reliable low technology graphical calculation aids and high technology computers with touch screen interfaces offers potential for improvement in patient safety; however further development and stringent evaluation are required before deployment in safety critical environments.

I. Introduction

When we go into hospital we expect to be treated well; yet approximately 11% of patients in hospitals suffer adverse events, of which half are preventable, and about a third of which lead to moderate or greater disability or death [1].

Medication errors (i.e., calculation and administration of the correct drug quantities and rates) are a common and preventable cause of such adverse events. Medication errors affect 2–14% of all hospital inpatients, and 1–2% of all hospital inpatients are harmed as a result; 7,000 are killed every year in the US alone [2]. More than 1 in 6 medication errors involve miscalculation of doses, incorrect expression of units, or incorrect rates of administration; and there are many case reports detailing how these errors have directly resulted in patient harm [3][4]. The above figures probably represent an underestimate: due to the complex multi-factorial nature of drug dose calculations, many errors go unrecognized or unreported, since if the clinicians have mistakenly determined that a calculation is correct and proceeded to deliver a drug dose, they are unlikely to report a calculation error has occurred. Furthermore, if patient harm occurs, one has to trace the cause back to an incorrect dose in order to report it, but patients frequently have co-morbitities that provide alternative explanations for the adverse clinical outcome.

Conventional methods of performing drug dosage calculation are more complex and error-prone than is usually appreciated. Indeed, increasing the number of steps, complexity, speed, workload, or stress levels all increase the probability of error [5]. It follows that users, regardless of skill, are likely to make slips that affect the results they calculate. Having defined the problems and highlighted their impact, we compare alternative approaches.

Our explorations raise deep concerns about the culture and assumptions surrounding the methods, including IT, which are used to perform calculations in the clinical environment; and the contributions of poor design to drug dose calculation errors. It is likely that the ideas discussed in this paper generalize to other areas of healthcare such as radiotherapy, and to wider applications outside of healthcare such as navigation and finance. However, discussion of topics other than drug dose calculation are beyond the scope of the present paper. Furthermore, we assume the user is skilled (for example, a qualified clinician) and that the nature of the calculations and their importance is well-understood.

A. *Previous work*

In this paper we examine and evaluate alternative approaches to more dependable calculation, and discuss the cultural barriers to the uptake of improved approaches. Curiously, the literature on the usability of general purpose calculators has placed greater emphasis on user models and education than on dependability [6], [7], [8]. It appears that the effectiveness and reliability of calculators is taken for granted; for most users they are neither new nor interesting — even though modern natural (e.g., gesture based) user interfaces offer considerable improvements [9].

Our earlier work (e.g., [10], [11], [12]) has, perhaps uniquely, been very critical of conventional calculator user interfaces. We have introduced special purpose calculators for drug dose calculation [10], and have shown that such well-designed user interfaces can halve the incidence of 'out by ten' calculation errors [11]. We have also shown that

978-0-7695-5089-3/13 $26.00 © 2013 IEEE
DOI 10.1109/ICHI.2013.62

numerical keyboards have approximately double the error rate of up/down number entry, and that conventional calculators and spreadsheets have unacceptably high error rates, which are avoidable by improved design [13]. We therefore consider application of our previous findings to suggest alternative designs that have potential to improve accuracy and reduce error rates in the clinical environment.

II. A ROUTINE DRUG DOSE CALCULATION

The example used in this paper is taken from a real case that resulted in the death of a patient, with subsequent root cause analysis (RCA) that was reported by the Institute for Safe Medication Practices (ISMP) [4], [10], [14].

The cause was the sort of calculation some clinicians do every day: in chemotherapy, a patient is to be given 5,250 mg fluorouracil at a concentration of 45.57 mg per mL over 4 days. What is the rate in mL per hour to program their infusion pump?

A. Calculation is not the only problem

Although calculation is a critical part of delivering the right drug in the right quantity and rate to the right patient (and by the right route, etc.), it is only part of the task that should be scrutinized.

Information design is crucial: users need to know what calculation they are expected to perform. The ISMP report [4] criticizes the information design of the drug bag label shown in figure 1: there is far too much irrelevant information, and the relevant information is not presented in a way that makes the user's tasks (e.g., identifying the patient or calculating a rate in mL per hour) easier. The first line ends "m" as the original label was not long enough to print more; possibly "L)" has been omitted. Note that the label refers both to days and to hours (as units of 24h, despite ISMP rules suggesting that "h" is an error prone abbreviation that should be avoided, and instead should be written as "hr"). Since the patient can read this label, it may have been helpful to say, "Bag will last 4 days at full usage with 12 hours reserve," to provide a potential means of cross-checking, rather than "14.8 mL reserve," which in itself is not very useful information.

The social culture users operate in is crucial. When errors occur in healthcare, there is widespread tendency to blame users rather than the systems they use. There is also a very common view that new technology is better, so remaining errors must be due to use error. We will discuss healthcare culture further in section VI; in particular, since our study shows that an old, low technology solution is effective, we will discuss in depth some misconceptions about computing technology.

Finally, although this paper is self-contained, the Appendix explains important parts of any drug dose calculation task that a clinician must perform as well as the bare calculation. This is the central concern of the present paper: calculators should not be considered in the abstract, but in the full context of the associated requirements, tasks, and environment that they support.

CHEMOTHERAPY DISPOSE OF PROPERLY

```
FLOUROURACIL 50 mg/mL INJ 5924.48 mg (118.49 m
In D5W IV Total Volume: 130 mL
Final Concentration: 45.57 mg/mL
Dose: 5250 mg/4days (1312.5mg/24h)
Rate: 28.8mL/24h (1.2mL/h) Bag will last 4 days
at full usage with 14.8 mL reserve.
Dr. [redacted] Rx#ABS19073
Prep: JUL 31 2006  905 Exp: 7days
[redacted] Pharmacy [redacted]
11560 [redacted] Ave. [redacted]
```

Fig. 1. Reproduction of the original drug bag label based on a photograph from [4]. The figure accurately reproduces the text, including character spacing, line breaks and font; the text "ABS19073" and the "905" were written by hand — that S might be a badly-written 5; and the "JUL 31 2006" was rubber-stamped. The blackened ■ regions are obscured to preserve anonymity

III. FIVE CALCULATION CONDITIONS

We now explore five conditions of drug dose calculation: pen & paper, general purpose calculators, spreadsheets, dedicated calculators, and nomograms. Nomograms are less familiar, and we will explain these calculation devices in more detail.

A. Dose calculation using pen & paper

Doing the drug dose calculation (see above) on paper starts off like figure 2 — if you can remember the technicalities of how to do long division and multiplication! Even for skilled users, the sums shown worked out in figure 2 require lots of fiddly work, as well as remembering to include the part of the calculation dividing by 24 hours in a day, which was not one of the numbers written down in the original problem. The case reported in [4] had two nurses both omit the 24 hours from their calculation.

There are many ways to do long division (figure 2 is the way we do it), and it would not be surprising if clinicians doing this calculation argue about "the" right way to do it. Unfortunately it would be very easy to use a correct method and still get the wrong answer through some other slip.

Of course one should not just do this calculation, as it is important to think through what needs doing and how to *independently* check it, as well as actually checking it. In other words, an important part of the task and mental workload (particularly for safety critical applications such as drug dose calculations) is not just doing the calculation but also doing adequate checking (see this paper's Appendix).

This routine drug dose calculation is so complex to do using pen & paper that most people would use a calculator — the condition we review next.

B. Dose calculation using a calculator

The Casio HS-8V is a commonly used general purpose calculator. Because the sum is complex either you need to use paper for intermediate results, or you have to use the calculator's memory function (or augment the use of the calculator with mental calculation). The display of such devices often

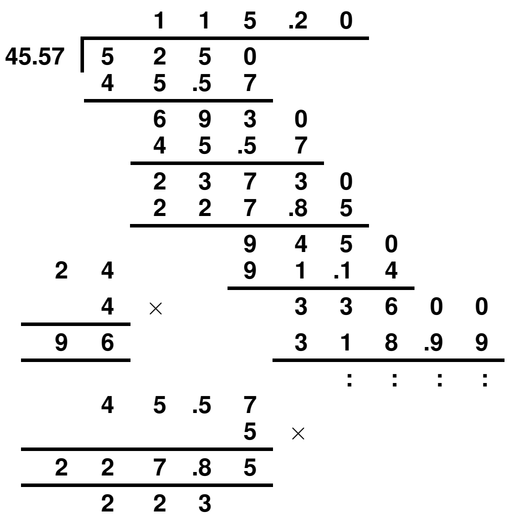


Fig. 2. Some of the paperwork for performing the calculations. Notice that decimals points may easily be misplaced or misread.

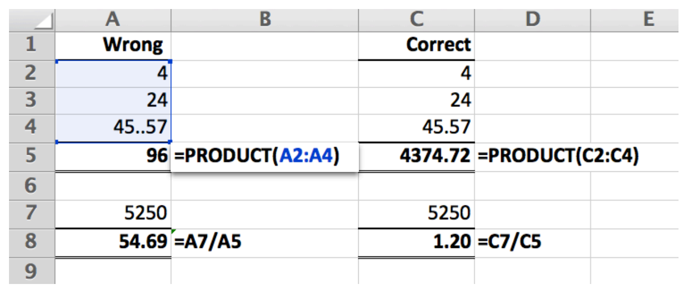
suffers from critically poor legibility, which further limits their appropriateness for clinical environments [12] (figure 3).

If memory is not zero, it is very hard to store numbers in the memory, since there is no "store in memory key" — instead, the user has to write the number down (subverting the whole point of memory) or do a complex sequence of keystrokes so that, despite the memory not being zero initially, the number can be stored. This sequence is so complex, a user would be advised never to rely on the calculator's memory on this or any similar calculator.

If we use the calculator's memory, the most efficient sequence of keystrokes (determined using a computer simulation with programmed state-space search) on this particular calculator requires 22 key presses:: AC MRC MRC 4 × 2 4 M+ AC 5 2 5 0 ÷ 4 5 • 5 7 ÷ MRC =. Other calculators will almost certainly need different sequences to get the right answer (e.g., the Apple iPhone calculator app is significantly different): the sequence used here may give an incorrect answer without warning if used on other calculators [15], [10]. We are therefore surprised that root cause analyses (e.g., [4]) do not specify the make and type of calculators involved in incidents. Furthermore, unlike a pen & paper method, most calculators leave no permanent record of the steps performed in a calculation, which makes identification and analysis of calculation errors impossible.

Ironically, in practice, using a calculator is probably less reliable than using a pen & paper; not least because we tend to trust calculators, even though different calculators work in completely different ways [16], [10]! Both pen & paper and calculators are *unavoidably* complex. Neither method tells you if you make a slip, and if you do make a slip, both methods will give an answer that is wrong. Both approaches need independent checking. Despite this, the majority of fluid and drug dose calculations in clinical areas are performed using general purpose calculators.

Fig. 3. The popular Casio HS-8V, a typical low-end handheld calculator with memory. The decimal point in the numeric display is almost unreadable, another usability problem with this type of calculator [12].



	A	B	C	D	E
1	Wrong		Correct		
2	4		4		
3	24		24		
4	45..57		45.57		
5	96	=PRODUCT(A2:A4)	4374.72	=PRODUCT(C2:C4)	
6					
7	5250		5250		
8	54.69	=A7/A5	1.20	=C7/C5	
9					

Fig. 4. The calculation, $5250/(4\times24\times45.57)$ performed twice in Microsoft Excel. The screen shot here shows the calculation done with a simple slip (column A) and correctly (column C). Excel does not tell us any error has happened, and it would be easy to read off the final dose rate of 54.69, which is 45 times too large. The screen shot has not been retouched.

C. Dose calculation using a spreadsheet

Now consider Microsoft Excel, which has been around for over 25 years and is regularly used for complex calculations, such as calculation of radiotherapy doses. A very simple calculation using our example of the drug dose problem (figure 4) demonstrates that a simple key press error (typing one dot extra) gives a wrong answer without reporting any error.

Apple's Numbers, an alternative to Excel, does the same things as Excel. Apple had the dilemma of copying Excel or doing it properly. They copied Excel. Clinicians are likely to use spreadsheets to perform more complex calculations (e.g., for radiotherapy doses); and spreadsheets are not routinely designed to cross-check results by performing calculations in different ways [11]. Ironically, the more complex the calculation, the more likely one is to make a slip. If you make a slip, the spreadsheet will give a wrong answer without warning — and the more complex the calculation, the harder it will be to spot a wrong answer.

D. Dose calculation using dedicated software

Bespoke software to calculate drug doses is readily available for a variety of platforms including computers and smart phones, and provides the benefits of rapid calculation and (not always!) error-checking; however modern software is so complex that it is not possible to test every possible combination of states for latent coding errors [17]. Software based solutions are therefore vulnerable to rare but potentially extreme errors. Most medical "apps" are currently unregulated, and it is not clear who is is legally responsible — the clinician or the software programmer — if patients come to harm as a result of coding errors. In view of these concerns, the Medicines and Healthcare Products Regulatory Agency (MHRA) in Europe and the Food and Drug Administration (FDA) in the USA have recently ruled that any software which performs fluid or drug dosage calculations for patient care purposes must be subject to the same rules of registration and validation as a medical device before it can be approved for use in these regions [18], [19].

Many software-based dedicated solutions include error-trapping routines to prevent errors of several orders of magnitude which result from accidental entry of data outside normal ranges. However the user interface of numeric data entry by a keypad remains vulnerable to unrecognized key-stroke errors of lesser magnitude, which have been estimated to occur in around 4% of key-presses [13]. Smart phones are popular and ubiquitous; however their use may be restricted in clinical areas due to the hospital policies for infection control and potential effects of electromagnetic radiation on other medical equipment (e.g., monitors, infusion pumps).

Unlike pen & paper calculations, most software-based solutions do not provide a hard copy of the data entered, details of steps in the calculation, and results. Hard copies are desirable for filing in the patient notes to form a permanent record: in the event of a drug dosing error they form an important legal record, and facilitate root cause analysis to precisely identify the source of error and prevent it happening again in future. Although some medical devices (e.g., syringe drivers) incorporate automatic keystroke logging to flash memory for this purpose, this is not the case for most software solutions and smartphone apps. Integrated "paperless" patient management IT solutions exist which store all patient-related data including medical notes, prescription charts and output from monitors for all aspects of the patient's admission to hospital. These cross-check drug doses and compatibilities against the patient's details, medical records, investigations, and physiological parameters; calculate appropriate drug doses with error capturing; and record drug administration. These systems are currently used at some centres in the US, but high capital and maintainance costs has meant that they are yet to gain widespread acceptance elsewhere.

E. Dose calculation using a nomogram

A nomogram or nomograph (Greek: $\nu\acute{o}\mu o\varsigma$ = "law" + $\gamma\rho\alpha\mu\mu\acute{\eta}$ = "line") is a graphical representation of a mathematical relationship that facilitates rapid repeated calculation. In its simplest form a nomogram consists of three scales drawn on a piece of paper (or a computer screen), one scale per variable, arranged and calibrated in such a way that a straight line connecting two of the independent variables intersects a third scale at a point corresponding to the solution. A nomogram with five scales is illustrated in figure 5.

It is worth noting a recent trend to misuse the word "nomogram" to describe regression analysis of morbidity and mortality data using a series of parallel weighted scales [20]: these are not nomograms as they do not perform a full calculation or express a mathematical relationship, and the user is required to perform the calculation in part by other means. The word nomogram is being used merely to mean "calculator" rather than using scales.

Nomograms have a long and distinguished history. Sir Isaac Newton used a combination of nomograms and slide rules to find roots of arbitrary polynomials [21], and Gerardus Mercator introduced his revolutionary world map with a graphical calculator (though using dividers rather than straight edges) as far back as 1569; but nomograms as we now know them (alignment charts with multiple axes) were first named and developed in earnest by Maurice d'Ocagne around 1885.

Since their modern inception in the 1880s, nomograms became widely used in the fields of engineering, astronomy, statistics, navigation and ballistics because of their convenience, reliability and usability for complex tasks. The advent of electronic computers and calculators in the 1960s saw their demise except in specialist areas where they are still used routinely for complex calculations where there are extreme time and accuracy constraints, as well as in environments where routine calculations must be performed conveniently without relying on electric power.

Ironically, the same computers that rendered nomograms obsolete are excellent tools for creating them, and this has resulted in a recent resurgence in interest in nomograms. Nomograms can solve complex mathematical problems [22], sometimes beautifully, and the computer tools available make it easy and pleasing to generate them. In the field of medicine, nomograms have been used to calculate physiological parameters [23], [24], [25], [26], [27], assess clinical risk [28], [29] and, more recently, for time-critical calculation of drug doses and fluid resuscitation regimes [30], [31], [32].

Figure 5 shows an original nomogram that can be used to work out the fluorouracil dose calculation in this article. The clinician solves the problem by drawing two lines on paper, thus making a permanent record of their calculation, as shown in the figure; the arrowed line shown solves the problem, and might even be drawn faintly (as in the figure) to be used as an example training pattern to help clinicians use the nomogram correctly. Because the lines must necessarily be drawn within the numerical range shown on the nomogram, the nomogram automatically ensures calculations stay within range; in the example here, it is plainly impossible to prescribe a total volume in excess of 150 mL.

In contrast to algebraic methods of calculation which are *imperative*, nomograms are *declarative* — i.e., they readily allow reverse and *what if?* calculations, which allow the user to easily check data entry, explore effects of perturbation of one or more input variables, and develop a qualitative understanding of the model. At a cognitive level, the interaction of the user with the graphical scales may lead the user to understand the mathematical relationships in a different way from that possible with key pads, equations, tables, and function graphs. This quality of nomograms may be of particular didactic value, for example in the communication of risks and benefits of treatment [20]. Patients can directly visualize the relative effects of various factors on their prognosis; and since nomograms are very low cost, readily printed, and easy to use, they may also take copies of the nomogram home with them as a record of their consultation, and a tool to improve understanding of their illness.

The paper record offers another very powerful benefit in error reduction: it is not just the result of the calculation, it also provides a permanent visual representation of the input variables, and how the calculation was performed. As the clinician slips the nomogram into the patient record folder, there will be other ones there. A glance will confirm whether the new nomogram works the same way: lines shaped, say, like $\wedge$ on one nomogram but shaped like $\vee$ on the previous ones would be an obvious indication of discrepancies. Similar application of a simple visual human-system interface to summarize a mass of complex data and rapidly convey an error situation was developed by NASA for the flight instrument displays (vertical altitude and velocity indicator, *VAVI*) of aircraft and spacecraft [33].

Nomograms are ideal for use in difficult environments such as in developing countries, expeditions and military use, where infrastructure, resources and enviromental factors make other methods (especially computers) unreliable or unusable — nomograms work after they have been dropped! In other safety-critical fields such as aviation and diving, nomograms and tables are always used in parallel with computers to cross-check results and to provide a primary means of calculation in case of computer failure [34].

IV. Comparing Methods

It is obvious that all conventional approaches (paper, calculator, spreadsheet) are so error-prone that anybody would be unlikely to get complex calculations right every day. It is amazing so many skilled clinicians do something so complex faultlessly all the time. If, as has happened, a patient is given a dose 24 times too high and dies, should we blame a clinician, or blame the system that asks clinicians to do something no normal human can do reliably? Perhaps we should be impressed the dose given is *exactly* 24 times too high, which means the clinicians have managed to do the sum correctly using the right formula with *only* one slip!

If you make a slip you are unlikely to notice it and are therefore even less likely to report it. There are probably many daily problems like this that go unreported, thankfully because the patient came to no obvious harm (though they may have stayed in hospital longer).

Paper, calculators and spreadsheets do not detect errors, as they have no idea what a clinician is trying to do. All these methods can produce incorrect results without warning. In contrast, with nomograms, range of use errors are impossible, and no significant new errors are introduced — a view which is supported by our experimental evidence.

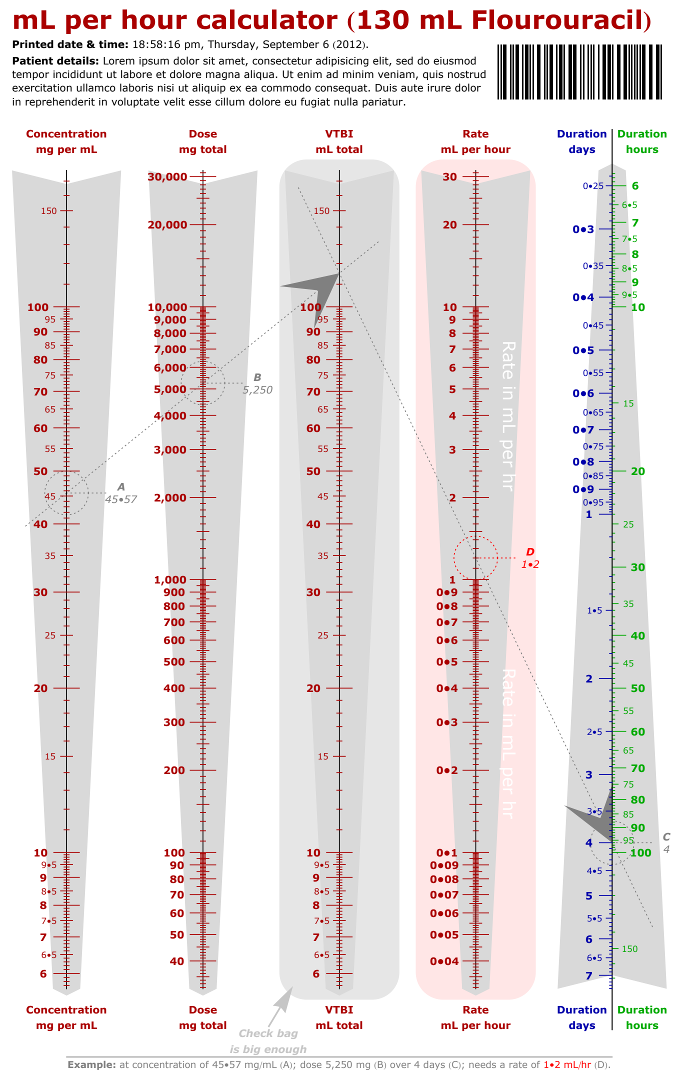


Fig. 5. A nomogram for this paper's drug dose calculation. The nomogram shown here is far too small to use at this scale, but a PDF of a full-sized version can be downloaded from `www.harold.thimbleby.net/nomogram`

Of course, the user still needs to check the patient, drugs, route and nomogram are properly related. That is why the nomogram shown has patient details on it and has a bar code to check against the patient's ID (wrist band and clinical notes). It can be printed on waterproof paper to make it more reliable for a hospital environment. The nomogram necessarily uses the right formula for the calculation, including remembering to divide by 24. The pharmacy computer that drew the nomogram sees to that.

One clinician can do the calculation on paper, and another can check the correct workins easily by checking the nomogram: it is just a matter of checking two lines are right and that the answer was used correctly, rather than checking lots of digits are right. File the permanent paper record of the

	Magnitude of error		
	Low	Medium	High
Calculator	0–25%	25–50%	50–75%
Nomogram	5.9%	1.2%	0%
Calculator	17.9%	14.3%	8.3%
Pen & paper	25%	16.7%	9.5%

	Likelihood ratios		
	0–25%	25–50%	50–75%
Calculator v nomogram	$p = 0.015$	$p = 0.001$	$p = 0.002$
Calculator v pen & paper	$p = 0.258$	$p = 0.670$	$p = 0.787$
Nomogram v pen & paper	$p < 0.001$	$p < 0.001$	$p = 0.001$
Overall	$p = 0.002$	$p < 0.001$	$p = 0.002$

Table 1. Summary of results. See also figure 7.

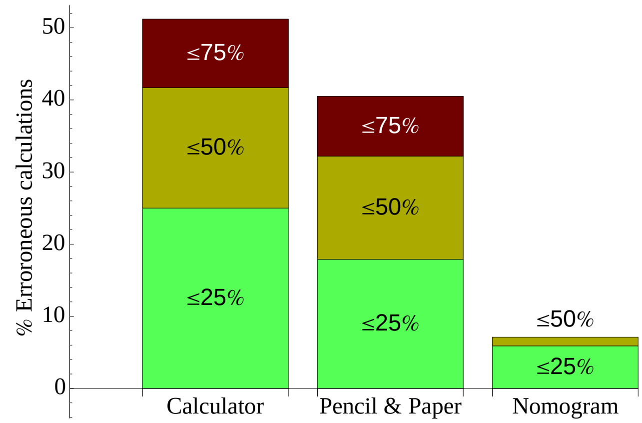


Fig. 6. Distribution of error magnitudes, visualising data from Table I. Thus 51.2% of calculator calculations have an error magnitude of 75% or less, compared to 7.1% of nomogram calculations having an error magnitude less than 75% and 50% — there are no errors in the range 50–75%.

calculation in the notes. Now just program the infusion pump to give 1.2 mL per hour …

V. Empirical evaluation of nomograms

Calculation of resuscitation fluid requirements in burns injuries involves a complex multi-step procedure that is prone to calculation errors, and involves application of a formula to calculate fluid requirements based on the patient's body weight and percentage body surface area burned (the "Parkland" formula is most commonly used worldwide [35]). This volume of fluid must then be given over a series of periods of different duration. In each case the rate of fluid administration must be calculated. Adjustments to the calculations must be made to take into account the time interval between the time of burn injury and the start of treatment; corrections for additional fluids given prior to hospital admission; and ongoing requirements for additional maintenance fluids. These calculations must be made rapidly under conditions of high task load and emotional stress, which further increases the likelihood that calculation errors occur.

Figure 7 and table I summarize findings from one of a series of randomized volunteer participant studies that compared the speed and accuracy of calculation methods to calculate fluid resuscitation requirements in both adult and paediatric burns in which 28 participants performed a total of 252 calculations on a series of computer generated simulated patient data sets using three different techniques: general purpose calculator, pen & paper, and nomogram. Computation speed using the nomogram was similar to that achieved with the calculator, and twice as fast as the pen & paper method (mean(SD): 73(31), 94(34), 214(103) s for calculator, nomogram, and pen & paper respectively). The nomogram produced fewer errors in all categories than the other two methods; the errors that did occur when the nomogram was used were of smaller magnitude, and therefore of lesser clinical significance than the other two methods; and the nomogram was deemed easiest to use by the majority of participants [16], [36].

Other studies using Bland-Altman analysis to compare the accuracy of nomograms and general purpose calculators in the calculation of drug doses and physiological variables show that the nomogram is significantly more accurate than a general purpose calculator [29], [30], [27]. Empirical experiments are not the only form of evaluation; for example Ed Hutchins presents a personal account emphasizing how easy nomograms are to use in navigation under adverse circumstances [37].

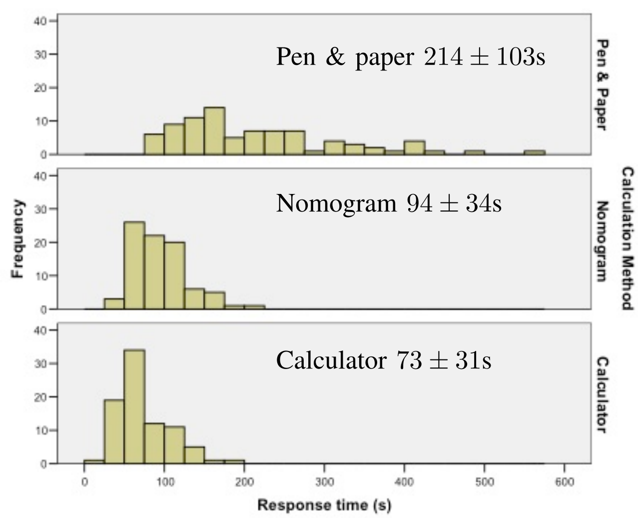


Fig. 7. Distribution of response times (seconds). Mean±standard deviation response times ($p < 0.001$) shown for each condition.

VI. Challenges to Healthcare IT culture

Nomograms are cheap, simple and effective; and they could save lives. One reason not to get too excited was well-put by Atul Gawande in his excellent book *The Checklist Manifesto* [38], where he writes about the WHO Surgical Checklist — another cheap bit of paper — that reduces morbidity and mortality in surgery. If something good is free, it is often considered to be worthless rather than priceless. Nobody is able to make a profit out of just printing bits of paper. What modern pharmacy would think it an improvement using nomograms, an idea familiar from the seventeenth century, when they could be buying modern IT? Who would want to consider anything other than computers when there are financial incentives to use them and financial penalties for not using them [39]?

However it is exciting to see the development of interactive nomograms, for instance on the iPad [40] and on the Interactive Nomogram Creation Tool web site [41]. Interactive nomograms could be further developed *and evaluated* for use on infusion pumps and other devices with touch screens, thus avoiding transcription errors transferring numbers from the nomogram to the device. The infusion pump itself could generate and display the nomogram (perhaps in a compact graphic format such as a virtual slide rule [40]).

The drug label's QR code could specify the right nomogram for a handheld device to create; thus the QR code, in figure 8 will download the nomogram shown in figure 5 — and (if printed) one gets the benefits of paper (dependability, auditability, etc) as well as the benefits of IT (security, automatic checks, relation to patient records, etc).

Since each nomogram can be printed for a specific patient and dose, it can take account of patient weight, etc, so it builds in a "dose error reduction" range check; it would not be possible to calculate an overdose. With the nomogram here, an easy limit to enforce is that, as the bag is known to be 130 mL, the total dose (i.e., dose in mg/concentration in mg per mL) physically cannot exceed 130 mL, and the nomogram scale could end at 130 mL, or have a bar across it. Of course, one is led to wonder if such "dose error reduction" calculations can be made by the pharmacy, why not just print the correct answer (1.2 mL per hour) on the bag and not risk human error in its repeated calculation? Figure 8 gives an example. As well as highlighting the required dose, we have made the expiry, the times, etc, easier to read; we removed the 28.8 mL per day dose as we know the infusion pump on the ward has to be programmed in mL per hour; we removed the 14.8 mL reserve (who cares about the *volume* that is left in the bag?), and replaced it with the *time*, 12 hours, which might be more use for the patient to plan the timings of their next trip to the hospital. Depending on the therapy, one might choose to make different information more prominent; here, we made the dose rate in mL per hour prominent, as the incorrect calculation of this was a factor in the fluorouracil overdose fatality [4] mentioned above.

Figure 8 gives an example of how we could improve the information design of a drug label, but why did prevailing culture allow such a badly designed label to be used in the first place? Where is the evidence of best practice? Why does the design require clinicians to repeat a complex calculation that has already been done for them, but is disguised by extraneous and confusing information? The culture that wants to solve problems with "off the shelf" IT solutions risks addressing the wrong issues, and leaving the deeper ones untouched.

One author, Grimes, has heavily criticized nomograms, likening them to "an epidemic of obsolete practice" making exaggerated claims in favour of modern IT [42], [43]. However his work contains a number of elementary factual errors and assertions for which he been unable or unwilling to provide evidence when challenged [44], [20].

An important misunderstanding surrounds Grimes's assertion that nomograms are inaccurate. In fact nomograms are generally more *accurate* (give an answer closer to the true value) although less *precise* (provide fewer digits) than general purpose electronic calculators. Electronic calculators that provide answers to 8 or more decimal places create a *false* sense of reliable precision and security for the unwary!

CHEMOTHERAPY DISPOSE OF PROPERLY

```
FLOUROURACIL 50 mg/mL
Patient: XXXX
Dose: 5250 mg/4days (1312.5 mg/day)
INJ 5924.48 mg
(118.49 mL in D5W IV)
Total Volume: 130 mL
Final Concentration: 45.57 mg/mL
Bag will last 4 days at 1.2 mL per hour
   with 12 hr reserve
Dr. XXX    XX Rx#ABS19073
Prep: 31 July 2006 @ 9:05
Expiry: 7 August 2006
 XXXX  XX Pharmacy XX     XX
11560 XXX   X Ave. XX    XXX
```

1•2
mL per hr

Fig. 8. Mock up of an improved drug bag label, based on the information shown in figure 1. There are still many ways of further improving the information design, beyond the scope of the present paper. Critically, the ISO standards 9241, 62366, etc, require iterative design: any proposed label design should be evaluated with users performing the appropriate tasks and improved.

Nomograms typically provide a precision of about 3 significant figures, or less than 1% error, [22], which is more than sufficient for most medical applications. The precision of nomograms is limited by factors such as the visual acuity of the user, parallax error or interpolation error when reading the scale graduations, and the physical size of the chart (which requires a compromize between being large enough to offer high precision, yet small enough for practical convenience). Any reduction in precision is more than compensated for by speed of use, and decrease in the frequency and magnitude of errors compared to computerized methods [16], [36].

Grimes [42] believes "old" implies "obsolete" and that modern computing is better than lower-technology solutions; a view that is echoed by some other authors:

> "Computerized approaches are ideal [...] because reliability can approach 100%, while methods that rely on human inspection will always miss some errors." [45]

Sometimes this may be the case, but arguments for or against the use of graphical or IT solutions without appropriate evidence can only be a basis for uninformed and misleading speculation. The broader question is what are the safest forms of calculation for a given context. No system of calculation is perfect. Every method has different strengths and weaknesses, which are determined by a combination of intrinsic and extrinsic factors — i.e., the culture and environment in which these are embedded.

Healthcare, which is supposed to have adopted "evidence-based practice," is in some quarters blind to the need to evaluate IT. IT is not automatically good, it depends on issues including situated use, cultural practices, use error, and so forth. In the meantime, this widespread unquestioned support of IT means that when errors occur, there will be a misleading tendency to blame the clinician rather than the supposedly "reliable" IT.

A widely reported computerization in a hospital [46] doubled fatality rates in a paediatric ward. One reason was that, effectively, every patient lost a clinician because they now had to work on a computer rather than be hands on. More generally, any IT intervention has to provide benefits that are at least equal to the effects of the loss of clinical staff time now turned to using the IT itself; the "obvious" benefit of the IT has to be balanced against this less obvious cost of using it (including managing any errors that arise during its use) — and this is a complex balance that at present requires empirical evaluation (indeed, as recommended by international standards such as ISO 62366).

A change in culture is required, and if properly implemented, there is every reason to believe that patient safety could be dramatically improved at little additional cost. In the meantime, there is a besetting problem — and a simple start at a solution.

The besetting problem is that "computerizing" what hospitals are currently doing will no doubt make them more efficient, but it does not address any underlying problems. People are approaching the wrong computational problems; calculators, Excel, conventional IT, clouds, are nothing but superficial help. Computers have to be used, and therefore their contribution has to at least exceed the consequences of the loss of staff time their use incurs [46]. It is glib to say it, but we'll just have adverse incidents faster, not less frequently or less harmful.

VII. Conclusions

Using a well-documented case study [4] as a starting point we have shown that (*i*) routine calculations can lead to adverse incidents; and (*ii*) appropriate use of technology can reduce error rates.

All systems of calculation have inherent advantages and limitations depending on context. Graphical methods of calculation have a number of unique advantages over electronic methods, and as such remain a valuable alternative primary means of calculation, as well as a means of rapidly cross-checking calculations performed by electronic devices. The possibility of combining elements from both methods offers exciting opportunities to design new types of human-computer interfaces that have potential to reduce the risks of calculation errors and patient harm.

What we can hope for is that one day in the future, if we are lucky, we will have dependable computers properly integrated into a more effective healthcare system; and that these improvements will be evidence-based. In the meantime, few want to put the research effort in to improving things when the hospitals themselves don't demand a better system, and when technophiles think it self-evident that computerization means progress — a point also made elsewhere [47]. To ignore the evidence and legal requirements (e.g., international standards such as ISO 62366) to follow human factors-informed user interface design is a dangerous and naïve mistake [48].

When an incident occurs, it is far too easy (and far cheaper) to blame the clinician who pressed "the wrong button". The ease of ignoring root causes, and root cause analyses often ignoring device-induced error, perpetuates the myths. We should be examining the unnecessary complexity, the unnecessary design faults, and, underlying it, the way we are failing to address the broader picture [46], [49], [14], [50]. If — when — we do that, the long range impact should be a reduction in the number of unnecessary hospital deaths and healthcare costs (and social costs), a rate that currently exceeds the death rate on the road, and deserves all the deep attention it can get [49].

Issues raised from this paper include:

- Newer technology does not necessarily improve usability.
- Nomograms are more accurate than and as fast as calculators (and pen & paper). However this does not necessarily mean they can achieve that performance in a clinical environment where *all* calculations are performed by nomograms. Clinicians might pick up the wrong nomograms; thus, one solution introduces a new form of error ("wrong nomogram") that previously was not possible. More work is clearly needed.
- Nomograms are easy to use, and they conform well to the larger clinical task — they can be personalized and made treatment-specific, they are visual (so they can be used and checked by more than one person), they can be designed to avoid out-of-range errors, and they leave a complete permanent record for clinical records.
- The classic HCI literature, particularly on calculators, does not address dependability and accuracy, nor does it really help appreciate the larger cultural issues (such as healthcare) in which calculators are embedded.
- Healthcare culture is a factor in sustaining suboptimal delivery of care. Computerising without careful (and iterative) user centred design may computerize suboptimal procedures. The benefits of embedding HCI experts (knowledgeable in both computer science and human factors) in healthcare should be evaluated.
- Calculators and apps are potentially hazardous and should be regulated like other medical devices.

Acknowledgements

With thanks to Russell Beale, Ann Blandford, Todd Johnson, Paolo Masci and Patrick Oladimeji. Partly supported by EPSRC Grant [EP/G059063/1].

References

[1] C. Vincent, G. Neale, and M. Woloshynowych, "Adverse events in British hospitals: Preliminary retrospective record review," *British Medical Journal*, vol. 322, pp. 517–519, 2001.

[2] D. Williams, "Medication errors," *Journal Royal College of Physicians*, vol. 37, pp. 343–346, 2007.

[3] T. Lesar, L. Briceland, and D. S. Stein, "Factors related to errors in medication prescribing," *Journal American Medical Association*, vol. 277, pp. 312–317, 1997.

[4] Institute for Safe Medication Practices, *Fluorouracil Incident Root Cause Analysis*. ISMP, 2007. [Online]. Available: www.ismp-canada.org

[5] S. E. McDowell, H. S. Ferner, and R. E. Ferner, "The pathophysiology of medication errors: How and where they arise," *British Journal of Clinical Pharmacology*, vol. 67, pp. 605–613, 2009.

[6] F. G. Halasz and T. P. Moran, "Mental models and problem solving in using a calculator," in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI83*, 1983, pp. 212–216.

[7] R. E. Mayer and P. Bayman, "Psychology of calculator languages: A framework for describing differences in users' knowledge," *Communications of the ACM*, vol. 24, no. 8, pp. 511–520, 1981.

[8] R. M. Young, "The machine inside the machine: Users' models of pocket calculators," *International Journal of Man-Machine Studies*, vol. 15, no. 1, pp. 51–85, 1981.

[9] W. Thimbleby, "A novel pen-based calculator and its evaluation," in *Proceedings third Nordic Conference on Human-Computer Interaction, NordiCHI'04*, 2004, pp. 445–448.

[10] H. Thimbleby, "Ignorance of interaction programming is killing people," *ACM Interactions*, pp. 52–57, September+October 2008.

[11] H. Thimbleby and P. Cairns, "Reducing number entry errors: Solving a widespread, serious problem," *Journal Royal Society Interface*, vol. 7, no. 51, pp. 1429–1439, 2010.

[12] H. Thimbleby, "Reasons to question seven segment displays," in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI2013*. ACM, 2013, pp. 1431–1440.

[13] P. Oladimeji, H. Thimbleby, and A. Cox, "Number entry interfaces and their effects on errors and number perception," in *Proceedings IFIP Conference on Human-Computer Interaction, Interact 2011*, vol. IV, 2011, pp. 178–185.

[14] H. Thimbleby, "Is IT a dangerous prescription?" *BCS Interfaces*, vol. 84, pp. 5–10, 2010.

[15] H.Thimbleby, "Calculators are needlessly bad," *International Journal of Human-Computer Studies*, vol. 52, no. 6, pp. 1031–1069, 2000.

[16] O. Bodger, A. Theron, and D. Williams, "Comparison of three techniques for calculation of the Parkland formula to aid fluid resuscitation in paediatric burns," *European Journal of Anaesthesiology*, vol. 30, pp. 1–9, 2013.

[17] H. Thimbleby, "Contributing to safety and due diligence in safety-critical interactive systems development," in *Proceedings ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS09*, G. Calvary, T. C. N. Graham, and P. Gray, Eds. ACM, 2009, pp. 221–230.

[18] A. J. Barton, "The regulation of mobile health applications," *BioMed Central Medicine*, vol. 10, pp. 46–49, 2012.

[19] (Retrieved June 15, 2013) Regulation of health apps: A practical guide. [Online]. Available: http://blog.d4.org.uk/2012/01/regulation-of-health-apps-a-practical-guide.html

[20] M. W. Kattan and J. Marasco, "What is a real nomogram?" *Seminars in Oncology*, vol. 37, no. 1, pp. 23–26, 2010.

[21] C. J. Sangwin, "Newton's polynomial solver," *Journal of The Oughtred Society*, vol. 11, no. 1, pp. 3–7, 2002.

[22] R. Doerfler, A. Gluchoff, S. Guthery, and P. Hamburg, *Calculating Curves: The Mathematics, History and Aesthetic Appeal of T. H. Gronwall's Nomographic Work*. Docent Press, 2012.

[23] W. Boothby and R. Sandiford, "Nomographic charts for the calculation of the metabolic rate by the gasometer method," *Boston Medical and Surgical Journal*, vol. 185, no. 12, pp. 337–354, 1921.

[24] M. Noda, "BMI nomogram [corrected]," *Clinical Nutrition*, vol. 27, no. 1, pp. 168–169, 2007.

[25] O. Siggaard-Andersen and K. Engel, "A new acid-base nomogram. an improved method for the calculation of acid-base data," *Scandinavian Journal of Clinical and Laboratory Investigation*, vol. 12, no. 2, pp. 177–186, 1960.

[26] M. Gulati, H. R. Black, L. J. Shaw, M. F. Arnsdorf, C. N. B. Merz, M. S. Lauer, T. H. Marwick, D. K. Pandey, R. H. Wicklund, and R. A. Thisted, "The prognostic value of a nomogram for exercise capacity in women," *The New England Journal of Medicine*, vol. 353, no. 3, pp. 468–475, 2005.

[27] J. Walker and D. Williams, "A nomogram to calculate estimated glomerular filtration rate (eGFR)," *European Journal of Anaesthesiology*, vol. 30, p. 152, 2013.

[28] N. D. Nguyen, S. A. Frost, J. R. Center, J. A. Eisman, and T. V. Nguyen, "Development of a nomogram for individualizing hip fracture risk in men and women," *Osteoporos International*, vol. 18, no. 8, pp. 1109–1117, 2007.

[29] D. Williams and J. Walker, "A nomogram to calculate physiological and operative severity score for the enumeration of mortality and morbidity (POSSUM)," *European Journal of Anaesthesiology*, vol. 30, pp. 234–235, 2013.

[30] D.Williams and J.Walker, "A nomogram to prevent local anaesthetic toxicity," *European Journal of Anaesthesiology*, vol. 30, p. 242, 2013.

[31] D. Williams and R. Doerfler, "Graphic aids for calculation of fluid resuscitation requirements in pediatric burns," *Annals of Plastic Surgery*, vol. 69, pp. 260–264, 2012.

[32] D. Williams, "Nomograms to aid fluid resuscitation in acute burns," *Burns*, vol. 37, pp. 543–545, 2011.

[33] C. A. Smith, M. Cummings, L. M. Forest, and L. J. Kessler, "Utilising ecological perception to support precision lunar landing," http://web.mit.edu/aeroastro/labs/halab/papers/HFES_758_Final.pdf, accessed: 12 June 2013.

[34] A. T. S. Report, *Aviation Research and Analysis Report AR-2009-052: Take-off Performance Calculation and Entry Errors: A Global Perspective*. Australian Transport Safety Bureau, 2009.

[35] S. Hettiaratchy and P. Papini, "Initial management of a major burn: II — Assessment and resuscitation," *British Medical Journal*, vol. 329, pp. 101–103, 2004.

[36] A. Theron, O. Bodger, and D. Williams, "Comparison of three techniques using the Parkland formula to aid fluid resuscitation in adult burns," *Emergency Medicine Journal*, vol. In Press, 2013.

[37] E. Hutchins, *Cognition in the Wild*. MIT Press, 1995.

[38] A. Gawande, *The Checklist Manifesto*. Profile Books, 2010.

[39] R. Koppel, S. M. Davidson, R. L. Wears, and C. A. Sinsky, "Health care information technology to the rescue," in *First Do Less Harm: Confronting the Inconvenient Problems of Patient Safety*. Cornell University Press, 2012, pp. 62–89.

[40] NomoGraphics Oy. (2011) Nomographics. [Online]. Available: www.nomographics.com

[41] T. Jones, R. Camara, and A. Marrero. (retrieved Sept 2012) Interactive nomogram creation tool. [Online]. Available: tinyurl.com/2h4jkf

[42] D. A. Grimes, "The nomogram epidemic: Resurgence of a medical relic," *Annals of Internal Medicine*, vol. 149, no. 4, pp. 273–275, 2008.

[43] D.A.Grimes, "Response: Where is the evidence that the resurgence of nomograms is harmful?" *Annals of Internal Medicine*, vol. 150, no. 4, p. 282, 2009.

[44] S. H. Zwillich, "Where is the evidence that the resurgence of nomograms is harmful?" *Annals of Internal Medicine*, vol. 150, p. 282, 2009.

[45] D. W. Bates, D. J. Cullen, N. Laird, L. A. Petersen, S. D. Small, D. Servi, G. Laffel, B. J. Sweitzer, B. F. Shea, R. Hallisey, M. V. Villet, R. Nemeskai, and L. Leape, "Incidence of adverse drug events and potential drug events," *Journal of the American Medical Association*, vol. 274, no. 1, pp. 29–34, 1995.

[46] Y. Y. Han, J. A. Carcillo, S. T. Venkataraman, R. S. Clark, R. S. Watson, T. C. Nguyen, H. Bayir, and R. A. Orr, "Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system," *Pediatrics*, vol. 116, pp. 1506–1512, 2005.

[47] A. Blandford, G. Buchanan, P. Curzon, D. Furniss, and H. Thimbleby, "Few are looking: Invisible problems with interactive medical devices," in *ACM Workshop on Interactive Systems in Healthcare (WISH), Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI2010*, 2010, pp. 9–12.

[48] T. Landauer, *The Trouble with Computers*. MIT Press, 1995.

[49] Institute of Medicine, *To Err is Human*. National Academies Press, 2000.

[50] H. Thimbleby, "Reaching to the stars with IT projects," *Public Service Review: UK Science and Technology*, vol. 5, pp. 20–21, 2012.

[51] G. Downie, *Calculating Drug Doses Safely: A Handbook for clinicians and Midwives*, 2nd ed. Churchill Livingstone, 2010.

[52] R. Lapham and H. Agar, *Drug Calculations for clinicians: A step-by-step approach*, 3rd ed. Hodder-Arnold, 2009.

APPENDIX

The article explores a mathematical problem: A patient is given 5,250 mg at a concentration of 45.57 mg per mL of the chemotherapy drug fluorouracil over 4 days; find the dose in mL per hour. This appendix makes clear some of the additional numerical thinking a clinician must do as part of the clinical task in addition to the bare calculation — though note that much of the work discussed in this appendix is obviated with nomograms. Note that the problem is not merely mathematical: the body of the paper compliments this appendix by emphasizing the user interface, cultural, training and human factors issues that are essential to address even before a mathematical problem can be identified.

A. Is the formula right?

First, see what the dimensions of the numbers are; here we have mg, mg per mL, 4 days, and mL per hour. The 4 days do not fit; so how do we change them to hours, which are what are going to be needed in the end?

$$4 \text{ days} \times 24 \text{ hours per day} = 96 \text{ hours.}$$

Notice in converting days to hours, we also had the dimensional multiplication: days × hours / day. The two "days" cancel, and leave us with hours, as wanted. So, all we need to do with the other numbers is find a way of cancelling things to end up with mL per hour. Thus 5,250 mg / 45.57 mg per mL simplifies to (5,250 mg / 45.57 mg) mL, just by moving the dimensions around using basic algebra. We've ended up with a division of mg, so we can do that immediately, perhaps on a calculator. So, we've got 115.2 mL. We want so-many mL per hour; so we now need to convert mL into mL per hour. But we've got 96 hours, so here we go: 115.2 mL divided by 96 hr = 1.2 mL per hr. That could be done on a calculator.

The full calculation we did was $(5,250/45.57)/(4 \times 24)$. Does it make sense? If the patient was supposed to have more florouracil, the dose rate should be higher. Indeed, the quantity of fluorouracil is "on the top" and increasing it will increase the result. If the concentration was less, the patient would need a faster dose; indeed, the concentration is "on the bottom," so decreasing it makes the answer larger. Finally, if the duration was shorter, the rate would need to be faster; indeed, making the 4 days shorter/smaller (since it is on the bottom) would make the answer larger. So all the numbers are in the right place for it to work the right way.

B. Is the calculation right?

Is the answer about right? It is hard to do the sum here, but we can make it easier. How about this: $(5,000/50)/(4 \times 25)$? Those are all close numbers, and are much easier to work with. 4×25 is 100, for instance. $5,000/50$ is the same as $500/5$, which is 100. Finally $100/100$ is 1. That is pretty close to what we got with a calculator, 1.2. So, we've got two independent reasons to believe we used the right formula (dimensions and the increase/decrease arguments), and so there are two reasons to believe the numerical answer: we used an exact calculation and we used an easy approximation, and they were very close. Any error that mattered would of course have given a very different result, warning that at least one of the sums was wrong.

C. Independent double check!

We should get somebody else to do the *whole* calculation again without knowing what we did, or even what we were thinking. Perhaps there was some assumption we got wrong. Maybe the patient is supposed to have $1,000$ mg, or perhaps we misread the label. Somebody working from our figures would just make the same mistake if we had misread the figures. They have to start again, be encouraged to use a different method, and only compare their final answer with us.

Safety-critical fields such as aviation and diving require calculations to be independantly double-checked using two different methods of calculation. These are typically a bespoke computer-based method (e.g., a flight computer or dive computer); and a bespoke graphic method (e.g., nomograms, tables, or circular slide rule / "whizz wheel") [34].

A weaker form of independent double check is to deliberately perform the calculation using the same method, but in a different way. For example, rearranging the calculation shows

$$(5,250/45.57)/(4 \times 24) = 5250/(4 \times 24 \times 45.57)$$

(as used in figure 4) so both calculations should be performed to check if they are equal.

We know that calculators and spreadsheets often treat decimal points incorrectly [11], and since the user may be making systematic keystroke slips performing a similar calculation to see if the answer is approximately the same is also helpful. For example, $(5,000/40)/(4 \times 20)$ is easier to key in and has no decimal points, and has the reassuringly close value 1.56.

D. Best practice

Note that the pharmacy can help a lot, by providing clear instructions, including details of the calculations — or a nomogram. One also wonders why the pharmacy didn't make the numbers easier to work with — like we did in the double-checking. For example, they could have tweaked the real concentration to exactly 50 mg per mL rather than leaving it at the more awkward 45.57 mg per mL (adjusting other values to compensate if necessary).

People need to learn from both correct and incorrect calculations. From correct calculations: what methods work, and can new users learn these methods? From incorrect calculations there is a need in general for users to learn what to avoid, and when things go wrong to see what errors were made to determine how to change systems and practice. Doing any of this requires the calculations and all relevant details (such as the date, the type of calculator used, the identities of the people doing and checking the calculation, etc.) need to be recorded. The formal root cause analysis [4], used as the concrete example in this paper, is notable in that it has virtually no information about how the problematic calculation was made.

This appendix raises issues that are all common sense, but for best clinical practice, please see the latest editions of [51], [52]. (These are UK books; other countries will have other procedures that should be considered.)