



University of Dundee

Reflex and reflective testing

Murphy, Michael J.

Published in:
Annals of Clinical Biochemistry

DOI:
[10.1177/0004563221993153](https://doi.org/10.1177/0004563221993153)

Publication date:
2021

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):
Murphy, M. J. (2021). Reflex and reflective testing: progress, but much still to be done. *Annals of Clinical Biochemistry*. <https://doi.org/10.1177/0004563221993153>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Annals of Clinical Biochemistry

Reflex and reflective testing: progress, but much still to be done

Journal:	<i>Annals of Clinical Biochemistry</i>
Manuscript ID	ACB-21-018
Manuscript Type:	Editorial
Date Submitted by the Author:	13-Jan-2021
Complete List of Authors:	Murphy, Michael; University of Dundee, Department of Biochemical Medicine
Keywords:	Laboratory management < Laboratory methods, Computers < Laboratory methods

SCHOLARONE™
Manuscripts

Author Accepted Manuscript

Reflex and reflective testing: progress, but much still to be done

Michael J Murphy

Department of Biochemical Medicine, Ninewells Hospital & Medical School, Dundee DD1 9SY, UK

Word count: 1224

Figures: 1

Author for correspondence:

Dr Michael Murphy, Blood Sciences, Level 7, Ninewells Hospital & Medical School, Dundee DD1 9SY, United Kingdom
T: 01382 383541
Email: m.j.murphy@dundee.ac.uk

DECLARATIONS

Conflict of interest: None

Funding: This editorial received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval: Not applicable

Guarantor: MM

Contributorship: MM sole author

Key words: Reflex testing, Reflective testing, Randomised controlled trial, Evidence-based medicine, Best practice

Clinical biochemists add value collectively by participating with other health professionals in the delivery of a high-quality service. They may also add value individually in various ways, including clinical liaison, adding comments and adding tests. Of these activities, the last is the most amenable to quantification; the metrics for adding value by adding tests are conceptually simple and readily applied. Reflex testing (the automatic addition of tests by analysers based on algorithms established by laboratory professionals) necessarily does most of the 'heavy lifting', reflecting the enormous throughput of heavily automated contemporary NHS laboratories. A much smaller number of tests is added reflectively by clinical biochemists. Paradoxically, reflective testing has received more attention, with early studies^{1,2} in particular seeking to prove the principle that it identifies patients who would otherwise be missed. However, subsequent studies^{3,4} have tended to report on both reflex and reflective testing. This is just as well, since they are indissolubly linked – and the impact of the choice of reflex threshold on the metrics of both has received inadequate attention.

The *Annals* has recently published a couple of papers which add significantly to what we know about the practice and value of adding tests. The first is a survey of reflex and reflective testing practice in the United

1
2
3 Kingdom (UK).⁵ This provides a detailed snapshot of current practice across a range of commonly encountered
4
5 diagnostic scenarios, and perhaps its most useful function is to set out the landscape for how tests are added.
6
7 Significant variation in practice was observed across a range of activities and scenarios. Examples include:
8
9 whether or not a test is added; if so, whether it is added reflectively or reflexly; whether interpretive
10
11 comments are included; whether the requestor is contacted before addition of tests. The results of the survey
12
13 thus provide a useful benchmarking function for laboratory professionals. They also provide a valuable first
14
15 step in laying the foundation for identifying best practice in this area; as the authors highlight, such advice
16
17 exists for interpretation⁶ even though as pointed out above this is less readily measured.

18
19
20
21 A few observations are warranted. First, the thresholds used to trigger reflex addition of tests vary widely, e.g.
22
23 the hypocalcaemic threshold to trigger magnesium measurement varied from 1.50 mmol/L up to 2.20 mmol/L.
24
25 Even allowing for differences in the nature, size and staffing of hospital laboratories, and populations served,
26
27 the extent of the observed variation invites scrutiny. Second, in the table which documents quantitative
28
29 aspects of the survey, reflective thresholds are listed alongside the reflex thresholds for each scenario, as if
30
31 equivalent. However, reflective 'thresholds' must be interpreted cautiously, for two reasons: (a) the *raison*
32
33 *d'être* of reflective testing is that it permits more complex information to be taken into consideration than can
34
35 readily be incorporated into reflex algorithms⁷ – thus any 'threshold' is of its nature 'softer' – if 'hard', it is
36
37 effectively a reflex threshold; (b) the reflex thresholds/boundaries used in each diagnostic scenario affect the
38
39 clinical utility of reflective testing.⁴ Third, addition of tests is widely accompanied by interpretive comments,
40
41 although it is not clear from the results of the survey how these were broken down by reflex/reflective testing,
42
43 or where these were automated. In general, one might expect automated comments to accompany reflexly
44
45 added tests, and non-automated comments to be used to explain the more complex rationale of reflective
46
47 testing. In both cases, the addition of tests and comments is closely entwined. Finally, notwithstanding ethical
48
49 issues in the addition of some tests, e.g. diagnosis of malignancy, or pregnancy, it is clear that different
50
51 approaches are used in the binary decision to contact the requestor before adding the test, or not.

52
53
54
55
56 This issue sees the print publication of a second important paper in this area: a randomised controlled trial of
57
58 reflective testing in primary care patients.⁸ In this study, patients were randomly allocated to an intervention
59
60 group, where requesting clinicians received reflectively added test results and interpretive comments as

1
2
3 appropriate, in addition to the originally requested tests, and a control group where they did not. The medical
4 records of patients in each group were followed up six months after the reports/comments were issued.
5
6 Primary outcome measures were the adequacy or otherwise of intended and actual actions, as judged by a
7
8 multidisciplinary panel. Reflective testing was judged to be useful in 84% of cases, and favourably shifted the
9
10 distribution of adequate/neutral/inadequate actions.
11
12
13
14

15 As the first randomised controlled trial (RCT) of reflective testing, this study is a landmark step towards
16
17 providing an evidence base for adding tests. The outcomes were appropriately nuanced, accommodating a
18
19 spectrum of adequacy, and the elapsed interval post test/comment enough to remove any possibility of
20
21 confounding due to premature evaluation. The adjudicating panel was broadly based, including physicians
22
23 from both primary and secondary care, as well as a clinical chemist. The delta checks and thresholds are
24
25 supplied in supplementary material. In short this was a well-designed and executed study. A significant
26
27 omission was the failure to document the efficiency and effectiveness of the reflective testing in each group. In
28
29 a study where the other outcomes were 'softer' or at least more subjective, this would have provided helpful
30
31 quantitative detail.
32
33

34 This was, strictly, not an RCT of reflective testing; rather it was an RCT of reflective testing and associated
35
36 interpretive comments. As highlighted above and elsewhere,⁹ it is much harder to quantitate the value of an
37
38 interpretive comment than an added test; evaluating the combined effect of both activities 'muddies the
39
40 methodological waters'. However, since these activities are so closely related, some may judge that it is
41
42 appropriate to evaluate them together. Second, the finding that reflective testing was judged to be useful, and
43
44 added value, was not surprising. It would have been astonishing, and deeply troubling, if this were not the
45
46 case. Third, as the authors acknowledge, there was a significant – and unexplained – difference between
47
48 intervention and control groups in terms of consenting to participate, allowing for the possibility of selection
49
50 bias. Finally, the reports selected for potential reflective testing were pre filtered by computer
51
52 algorithms/delta checks, allowing for a different kind of selection bias.
53
54
55

56 What needs to be done next? The most pressing issue is to establish the variation in efficiency and
57
58 effectiveness across the range of reflex thresholds used in each diagnostic scenario. The variation in these
59
60 metrics by threshold has been established for TSH triggering free thyroxine measurement,⁴ but a

1
2
3 comprehensive evaluation is required covering other commonly encountered diagnostic scenarios. (Essentially
4
5 the data shown in Figure 1 for free thyroxine triggered by TSH needs to be replicated for other
6
7 triggering/triggered tests). The work involved is not trivial, but this is the first step to establishing optimal
8
9 thresholds to apply. It cannot be assumed that these metrics will vary in the same way for different scenarios.
10
11 The process of generating these data will reinforce awareness of the effect of reflex thresholds on the metrics
12
13 of both reflex and reflective testing. It will also provide the basis for establishing best practice in the addition
14
15 of tests. Establishing 'best practice' in the absence of data carries the risk of 'fossilising' activities which are not
16
17 evidence-based (laboratory professionals would require compelling reasons not to adhere to guidance issued
18
19 by national professional bodies). Reflective testing in particular is a discretionary activity and it is important
20
21 that we retain individual professional autonomy.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Legend to Figure 1

Number of diagnoses (cumulative) and numbers needed to diagnose (NND) for (a) hypothyroidism and (b) hyperthyroidism. Dotted lines indicate reflex thresholds. Reflective testing is confined to shaded areas. TSH, thyroid stimulating hormone. Reproduced with permission from reference 9.

References

1. Sinclair D, Spedding A, Young R. Can the laboratory affect the investigation and diagnosis of primary biliary cirrhosis? *J Clin Pathol* 2006;**59**:360-2
2. Preiss D, Todd L, Panarelli M. Diagnosing unsuspected hypopituitarism in adults from suggestive thyroid function test results. *Ann Clin Biochem* 2008;**45**:70-5
3. Jones BJ, Twomey PJ. Comparison of reflective and reflex testing for hypomagnesaemia in severe hypokalaemia. *J Clin Pathol* 2009;**62**:816-9
4. Srivastava R, Bartlett WA, Kennedy IM, Hiney A, Fletcher C, Murphy MJ. Reflex and reflective testing: efficiency and effectiveness of adding on laboratory tests. *Ann Clin Biochem* 2010;**47**:223-7
5. McKeeman GC, Hall SL, Freedman DB. Reflex and reflective testing practice in clinical biochemistry in the United Kingdom – a national survey. *Ann Clin Biochem* 2020;**57**:77-87
6. Kilpatrick ES. *Best practice when providing interpretive comments on laboratory medicine reports*. London: The Association for Clinical Biochemistry and Laboratory Medicine. September 2014

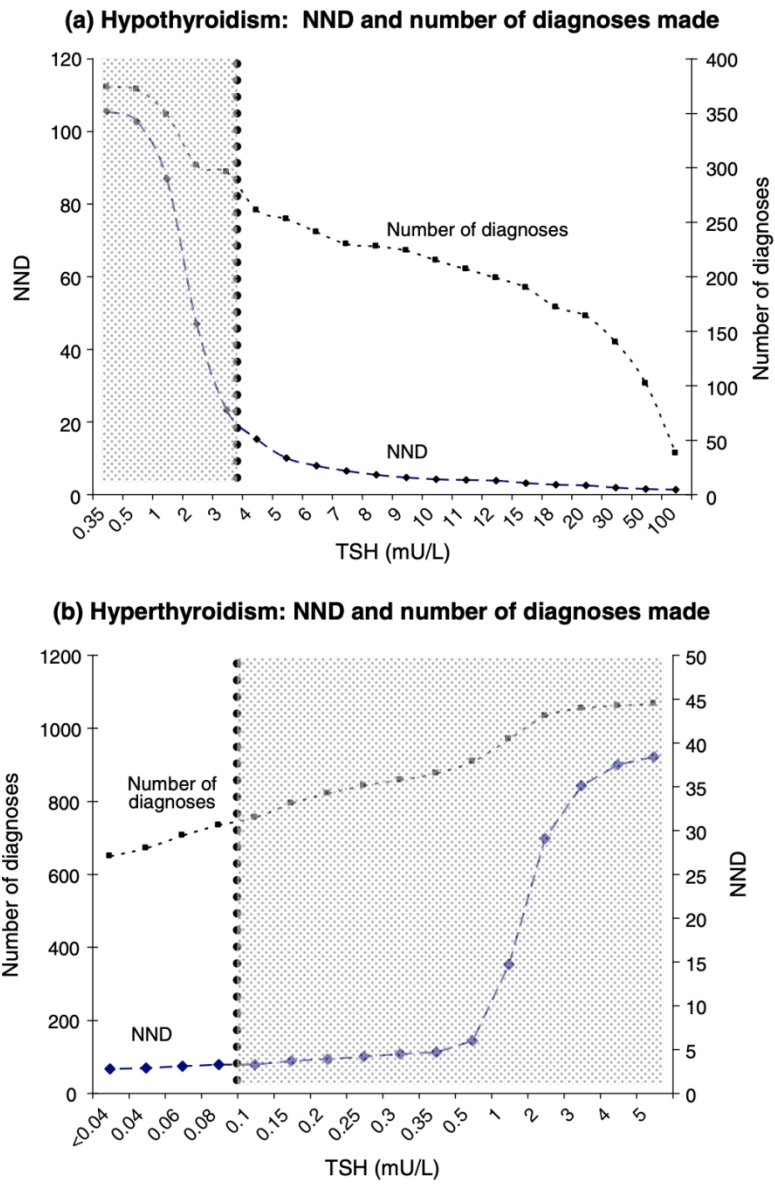
Author Accepted Manuscript

1
2
3
4
5 7. Paterson JR, Paterson R. Reflective testing: how useful is the practice of adding on tests by laboratory
6 clinicians? *J Clin Pathol* 2004;**57**:273-5
7

8
9
10
11 8. Oosterhuis WP, Verboeket-van de Venne WPHG, van Deursen CTBM, Stoffers HEJH, van Acker, BAG, Bossuyt
12 PMM. Reflective testing – a randomised controlled trial in primary care patients. *Ann Clin Biochem*
13 2021;**58**:xxx-x
14
15

16
17
18
19 9. Murphy MJ. Reflections on reflex thresholds. *Ann Clin Biochem* 2012;**49**:515-7
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Number of diagnoses (cumulative) and numbers needed to diagnose (NND) for (a) hypothyroidism and (b) hyperthyroidism. Dotted lines indicate reflex thresholds. Reflective testing is confined to shaded areas. TSH, thyroid stimulating hormone. Reproduced with permission from reference 9.