Published in final edited form as:

Pediatr Transplant. 2010 December ; 14(8): 940–943. doi:10.1111/j.1399-3046.2010.01396.x.

Is calculating the standard deviation of tacrolimus blood levels the new gold standard for evaluating nonadherence to medications in transplant recipients?

Eyal Shemesh, M.D

Department of Pediatrics and Kravis Children's Hospital, Mount Sinai Medical Center, New York

Richard N Fine, M.D.

Abstract

No it is not. Not yet. But looking at blood level variation might – in the near future - indeed become the de-facto gold standard for assessment of nonadherence in the transplant setting. If it does, this measurement method could lead to important discoveries not only for the transplant field but for adherence and outcomes research in other areas of medicine as well. Since nonadherence accounts for a majority of late rejections in pediatric solid organ transplant recipients (1–3), the discovery of a useful measure of adherence could be of great clinical importance in the provision of posttransplant care.

Brief history

The use of blood levels to determine adherence to medications is not new. A patient who is not taking a medication will have an undetectable level, a fact that is routinely utilized in medical practice. However, interpretation of a single level may be misleading and does not give a true picture of the long-term pattern of medication ingestion by a given patient. In addition, not taking a medication such as tacrolimus only once, or taking it only occasionally, would be expected to result in a blood level that is detectable but potentially sub therapeutic and thereby associated with a heightened rejection risk. Thus, it is important to look at a set of blood levels and not just at one level, and at nuances that go beyond simply looking at a level of zero or not, over time.

About a decade ago, a group of investigators at Mount Sinai Medical Center, supported by a Federal grant, first described the use of a novel measure of adherence to tacrolimus in children who had a liver transplant (4). The measure consisted of calculating the degree of fluctuation between individual measures of tacrolimus blood levels over a period of 1 year. This was done by calculating the Standard Deviation (SD) of each set of measures per patient. The higher the SD, the more extreme the fluctuation between individual measures. This suggests a pattern of medication ingestion which is erratic and likely to be the pattern in typical nonadherence. The idea of looking at variation of blood levels as a way of measuring the consistency of a medication regimen was later independently pursued by investigators from Cincinnati (5). However, the Cincinnati investigators originally targeted fluctuation in blood levels as though it is primarily determined by prescriber behavior (the physician's behavior) rather than by the patient's adherence behavior.

It became clear that indeed a high degree of fluctuation in blood levels predicts organ rejection episodes. Also, the Cincinnati investigators noted that an effort to standardize

prescribing behavior, which reduced some of the fluctuation in tacrolimus levels, did not limit rejection episodes. Further analysis revealed that rejection episodes in the Cincinnati cohort were primarily in patients with a very high level of fluctuation (SD of more than 2.0), who were “resistant” to efforts to standardize prescriber behavior (2). In other words, beyond a certain “threshold” of fluctuation, the fluctuation is related to patient characteristics (likely, adherence) and not to prescription patterns. In the decade since, studies of seven different pediatric transplant recipient cohorts (1–9) utilized this measure. Up to this point, data from several centers (Mount Sinai Medical Center in New York, UCLA in California, Cincinnati Children's Medical Center in Ohio) confirmed that an above-threshold level of SD of tacrolimus levels is a powerful predictor of late organ rejection, probably surpassing any other known such predictor (1–3,5). Now, additional data from the Hospital for Sick Children (SickKids), Toronto (in this issue) confirm and extend those findings. The SickKids manuscript by Pollock-BarZiv et al. extends and strengthens previous published results because it investigates several transplant populations (heart, kidney, and lung) in the largest study to date that studied this association.

Existing evidence

At this point, independent single-center studies of pediatric transplant recipients from 4 centers, in the US and Canada, involving in the aggregate several hundred patients followed for one to two years each, clearly establish the following: when a threshold value of 2 or more is applied, the SD measure is a strong predictor (predicts about 90%) of late rejection in patients with a variety of solid organ transplants. This measure is highly correlated with electronic monitoring of adherence (7), but electronic monitors are much harder to use (7,10). The SD value is correlated with some psychosocial constructs, notably posttraumatic stress related to transplantation (4–7), transition of care to adult providers (8), and missing procedures or clinic appointments (this issue). This method is easy to implement as it utilizes information routinely obtained in clinical practice (1). In a pilot study, a low-intensity behavioral intervention seemed to substantially reduce SD levels and reduce rejection rates in the studied cohort (11). In another pilot study, the use of cell phone reminders seemed to improve SD outcomes in the studied cohort (9). Thus, it is clear now that this measure is a strong predictor of late rejection, is related to adherence and psychosocial constructs such as distress and forgetfulness, and is sensitive to changes in those psychosocial constructs (psychosocial interventions can improve the SD numbers).

What do we still need to know?

And yet, questions still exist about the use of this measure. While practitioners seem to be willing to embrace this method (as it is predictive of outcomes and requires virtually no additional effort to measure), concerns voiced by adherence experts are summarized and explained below.

- A. All studies to date have been conducted in single centers. The relevant fluctuation threshold reported by each center varies somewhat although not much (from 2.0 to 3.5 SD units). Attempts to stratify results by age are largely lacking. It is argued that a multisite trial with numbers that are robust enough to try to stratify the results should be conducted before the measure is incorporated into practice.
- B. All previous studies were either retrospective or cross-sectional; a prospective study is desirable.
- C. Change in blood levels can happen as a result of factors other than patient adherence, for example absorption problems or change in prescriber practices. While this is correct, it is, in our review, a largely irrelevant concern. First, since this measure is looking at within-patient levels, the only absorption problems that

would be relevant here are those that are not stable within the same patient (i.e., worsening or transient). Those are unusual and, if they occur, are largely evident to the clinical team. Second, while prescriber practice could cause some increase in fluctuation, it is not an important reason for the clinically-important range of fluctuation, as demonstrated by Venkat et al. (2). Third, even if some unusual factor other than nonadherence could indeed move a few patients into the clinically worrisome range, the measure remains useful. Surely, a more thorough evaluation of the reason for fluctuation would be expected to unveil that hypothetical factor and may help in improving patient outcomes.

In our view, the use of this measure as a preferred measure of screening for nonadherence in the transplant population – at this point – is mainly subject to the conduct of a multisite, prospective, robust study that would look at the threshold across centers in relation to age and perhaps a few other disease variables. This task is currently under way as will be described below.

What are the ideal characteristics of a practical measure of adherence?

In our view, a practical measure of adherence should:

- A. Be easily integrated into practice. This would imply that the method does not use intensive or costly procedures over and above procedures that are routinely used clinically.
- B. Pose minimal or no additional burden to patients. The more burdensome a measure, the less likely are patients to follow it. Burdensome or intensive measurement activities (i.e., a long questionnaire, a burdensome electronic device, and so on) are almost guaranteed to lead to exclusion of the very patients that we are most interested in measuring (nonadherent patients).
- C. Be a direct measure of ingesting the medication, rather than an indirect measure. In particular, the measure should not rely on patient, parent, or provider self-report, as self reports have been shown to be inaccurate in prior research (i.e., 1).
- D. Rely on as simple a procedure as possible. This is because the more complicated the procedure (or calculation), the more opportunity there is for error. For example, while some reviewers advocated the use of a combination of measures (12), we believe that this, in fact, is not desirable. Adherence measures have repeatedly been shown, at times, to be uncorrelated with each other (i.e., 1,13), and thus “putting them together” might be nonsensical. In addition, we are not aware of any definitive, hypothesis-driven, prospective study that found that a combination of measures substantially improves the predictive validity of an existing, robust adherence assessment. Thus the use of a combination of measures does not have a strong empirical support at this point. It is also much harder to implement clinically. Hence it seems intuitive to conclude that a single measure, if robust enough, is preferred to a more burdensome combination technique.
- E. The measure should be correlated with adverse clinical outcomes, and the threshold at which those outcomes are present should be known. For clinical practice, we should be able to tolerate some trivial “nonadherence” (i.e., the patient is taking the medication an hour later on the weekends). We therefore should ideally know exactly at which point nonadherence becomes dangerous, in order to guide intervention efforts.
- F. The measure should be good enough to raise a valid concern about a patient's behavior, to be followed by more intensive queries. But it does not have to be fully

accurate in all cases. A quest for ultimate accuracy, in our view, brings one farther from discovering a clinically useful measure. We see no issue with accepting a measure of adherence that has a margin of error, especially if the measure offers substantial implementation benefits. Insisting on a cumbersome measure, even if it is more accurate, is the wrong approach in our view.

- G. The measure should provide evidence of nonadherence prospectively and in real-time, not only post-hoc, so as to enable efforts to remedy the problem before damage has already become irreversible.

The SD measure could conceivably meet all of those criteria. With regards to the last criterion, however, much more needs to be done. Most of the studies using a calculation of standard deviation of tacrolimus blood levels provide evidence of nonadherence after it has occurred over a protracted period of time. By then, clinical consequences, either subtle or clinically apparent, may have already occurred. We are aware of only one pilot study that used the SD data prospectively in an effort to prevent future rejections (11), with encouraging results (rejection rates dropped substantially). But much still needs to be done to prove that indeed SD values can be used proactively. Ideally, values would be calculated continuously and patient nonadherence would therefore be “captured” immediately when the threshold has been crossed; this approach has not yet been evaluated, except in the one pilot study mentioned above.

If it is proven to be accurate in predicting poor outcomes in a multisite setting, the SD method can be easily implemented as a clinical routine across transplant centers. Then, prospective intervention studies could move the field from the detection of nonadherence to improving outcomes. If alternative adherence assessment methods are discovered in the meanwhile that are as accurate and easy to implement, we might be in the very fortunate situation of having to choose between those methods, or being able to use different methods for different patients. At this point, however, in our view no other method of detection of nonadherence in the transplant setting has the same level of empirical support. In addition, no other method even comes close to having the same promising characteristics that would make it possible to implement on a large-scale, clinical level.

The next steps

We agree with Pollock-BarZiv et al (this issue) that “prospective studies are urgently needed”. Recently, the National Institute of Diabetes and Digestive and Kidney diseases (NIDDK) funded a multisite, prospective study to look at this measure over time in 5 transplant centers (ClinicalTrials.GOV identifier, NCT01154075). We hope and believe that this study will provide definitive information on the use of this measure.

Should we discover that this measure indeed predicts adverse outcomes with a given threshold, clinical implementation, we believe, could (and, perhaps, should) become the norm. Ample evidence already suggest that transplant centers are able to incorporate a measure of fluctuation in blood levels into their routine practice, and to act on this information to improve adherence (i.e.,11,14). Incorporating a measure of adherence into routine practice and following adherence behavior overtime is a goal that has rarely if ever been achieved in any medical discipline. Improving nonadherence seems possible once it is identified (9,11). Thus, the routine identification of nonadherence will provide the necessary first step towards improving posttransplant outcomes and survival. It will also be an unprecedented opportunity to research intervention methods to improve adherence: if robust adherence outcomes are monitored during routine clinical care, intervention research will benefit from a readily available outcome measure. In addition, patients could be stratified according to baseline adherence for differential intervention efforts – which is almost never

possible today. Interventions examined in this setting could later be implemented in other areas of medicine as well. Thus, the transplant community can become a de-facto “laboratory” that will examine innovative methods to improve adherence.

Ultimately, a reproducible marker of adequacy of immunosuppression, similar to the HgbA1c measure used for adequacy of diabetic control, would be an ideal surrogate to monitoring adherence in pediatric solid organ transplant recipients. Ideally, it will be possible to use a robust health outcome indicator in tandem with a behavioral marker (the SD measure). Until such a parameter is delineated, the calculation of the SD of tacrolimus blood levels could be used as a viable surrogate to evaluate the adequacy of immunosuppression as well as adherence in a given patient.

Conclusion

Looking at medication level variation to evaluate adherence in transplant settings has matured from an interesting idea to a very promising procedure. At this point, it is the most rigorously studied measure of adherence in pediatric transplant settings. However, its wide-scale implementation should be delayed until data are available from larger, prospective studies. If those become available, and the method is clinically implemented and is tied to prospective intervention efforts, it is likely that transplant programs will gain a substantial survival benefit from this discovery. It is also possible that the entire area of adherence research will see significant gains. We need to wait, but, hopefully, the wait will not be too long.

Acknowledgments

Dr. Shemesh's work on posttransplant adherence is supported by grant # R01 DK080740-02 from the National Institutes of Health / NIDDK.

REFERENCES

1. Shemesh E, Shneider BL, Savitzky JK, Arnott L, Gondolesi GE, Krieger NR, Kerkar N, Magid MS, Stuber ML, Schmeidler J, Yehuda R, Emre S. Medication adherence in pediatric and adolescent liver transplant recipients. *Pediatrics*. Apr.2004 113:825–32. [PubMed: 15060234]
2. Venkat VL, Nick TG, Wang Y, Bucuvalas JB. An objective measure to identify pediatric liver transplant recipients at risk for late allograft rejection related to non-adherence. *Pediatr Transplant*. Feb.2008 12:67–72. [PubMed: 18186891]
3. Stuber ML, Shemesh E, Seacord D, Washington J 3rd, Hellemann G, McDiarmid S. Evaluating non-adherence to immunosuppressant medications in pediatric liver transplant recipients. *Pediatr Transplant*. May.2008 12:284–8. [PubMed: 18331387]
4. Shemesh E, Lurie S, Stuber ML, Emre S, Patel Y, Vohra P, Aromando M, Shneider BL. A pilot study of posttraumatic stress and nonadherence in pediatric liver transplant recipients. *Pediatrics*. Feb.2000 105:E29. [PubMed: 10654989]
5. Bucuvalas JC, Ryckman FC, Arya G, Brandy A, Lesko A, Cole CR, Brent J, Kotagal U. A novel approach to managing variation: outpatient therapeutic monitoring of calcineurin inhibitor blood levels in liver transplant recipients. *J Pediatr*. 2005; 146:744–50. [PubMed: 15973310]
6. Shemesh E, Annunziato RA, Yehuda R, Shneider BL, Newcorn JH, Hutson C, Cohen JA, Briere J, Gorman JM, Emre S. Childhood abuse, nonadherence, and medical outcome in pediatric liver transplant recipients. *Journal of the American Academy of Child and Adolescent Psychiatry*. Oct. 2007 46:1280–9. [PubMed: 17885569]
7. Kerkar N, Annunziato RA, Foley L, Schmeidler J, Rumbo C, Emre S, Shneider B, Shemesh E. Prospective analysis of nonadherence in autoimmune hepatitis: a common problem. *J Pediatr Gastroenterol Nutr*. Nov.2006 43:629–34. [PubMed: 17130740]

8. Annunziato RA, Parkar S, Dugan CA, Barsade S, Arnon R, Miloh T, Iyer K, Kerkar N, Shemesh E. Brief Report: Deficits in Health Care Management Skills Among Adolescent and Young Adult Liver Transplant Recipients Transitioning to Adult Care Settings. *J Pediatr Psychol*. Dec 7.2009
9. Miloh T, Annunziato R, Arnon R, Warshaw J, Parkar S, Suchy FJ, Iyer K, Kerkar N. Improved adherence and outcomes for pediatric liver transplant recipients by using text messaging. *Pediatrics*. Nov.2009 124:e844–50. [PubMed: 19822583]
10. Shellmer DA, Zelikovsky N. The challenges of using medication event monitoring technology with pediatric transplant patients. *Pediatr Transplant*. Jun.2007 11:422–8. [PubMed: 17493224]
11. Shemesh E, Annunziato RA, Shneider BL, Dugan CA, Warshaw J, Kerkar N, Emre S. Improving adherence to medications in pediatric liver transplant recipients. *Pediatr Transplant*. May.2008 12:316–23. [PubMed: 18435607]
12. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. Aug 4.2005 353:487–97. [PubMed: 16079372]
13. Walshe L, Saple DG, Mehta SH, Shah B, Bollinger RC, Gupta A. Physician estimate of antiretroviral adherence in India: poor correlation with patient self-report and viral load. *AIDS Patient Care STDS*. Mar.2010 24:189–95. [PubMed: 20214487]
14. Emre S. Posttraumatic stress disorder in posttransplant children: creating a clinical program to address their needs. *CNS Spectr*. Feb.2006 11:118, 120–6. [PubMed: 16520689]