GRAND VALLEY STATE UNIVERSITY
SABBATICAL LEAVE REQUEST APPLICATION FORM

1. **Name** Edward Aboufadel 2. **Date employed by GVSU** August 1995

3. **College** CLAS 4. **Unit** Mathematics

5. **Semester(s) and year of latest sabbatical leave** Fall 2000

6. **Dates of Leave of Absence as faculty since the latest sabbatical leave** none

7. **Sabbatical leave is being requested for this year and semester(s)**

   Fall 2011 **Winter** _____ **Academic Year** _____

8. **Sabbatical Title** Mathematical Analysis of Blood Glucose Data from Type-1 Diabetics

9. **Are grants, fellowships, or other financial support expected during the sabbatical leave? Specify.**
   In conjunction with this proposal, I am planning on applying for a small grant from the Juvenile Diabetes Research Foundation.

10. **Is acceptance of a sabbatical award dependent on your receiving some additional support? Specify.** No.

11. **Where will the sabbatical work be done?** In the Grand Rapids area.

12. **In accord with the Sabbatical Policy as stated in Chapter 4, 2.30.4D of the Administrative Manual (see also Chapter 3, Section 3.04.C.5.d, Chapter 3, Section 3.04.E.6,7 and Chapter 4, Section 4.02 item 2.30.4.D, Faculty Handbook), I agree to submit a written account of my activities and a financial statement by the end of the term following my leave. I will remain at Grand Valley State University for at least one year after completion of my leave regardless of the length of the sabbatical.**

_____

Signature of Applicant Date

This sabbatical request has been reviewed by:

_____(approval, disapproval, or approval as modified)
Unit Date

_____(approval, disapproval, or approval as modified)
Applicant's Unit Head Date

_____(approval, disapproval, or approval as modified)
Chairperson, College Personnel Committee Date

_____(approval, disapproval, or approval as modified)
Dean Date

_____(approval, disapproval, or approval as modified)
Provost Date

SABBATICAL APPLICATION FORM FOR PREPARATION OF SABBATICAL LEAVE REQUESTS

A. This application is to be submitted to the College Dean no later than November 1 of the academic year preceding the academic year of leave. Unit review must therefore be complete before November 1.

B. Please type using 12 point font (either single or double space) a maximum of 10 pages for the narrative. Supporting letters and documents may be placed in an appendix. Use the "Sabbatical Leave Request Application Form" as the cover page to your proposal. Your Dean's office will tell you how many copies are needed.

C. Using this form, fill in the expandable gray boxes (Numbers 1-8) with narrative text. Please follow the instructions given under each heading, and be sure to complete each field.

1. Descriptive title of your project.

Mathematical Analysis of Blood Glucose Data from Type-1 Diabetics

2. Goals and objectives.

Proposals must have a clear conceptual focus. Be certain that the conditions and criteria for sabbatical leave as stated in the Administrative Manual, chapter 4, section 2.30.4.D), have been addressed. Proposals should be specific about the academic purpose of the leave. In particular, a proposal should conform to some or all of the following:

a) a planned effort to retrain professionally (with approval of the chair);
b) development of new capabilities for teaching or research;
c) a synthesis or development of prior effort or experiences;
d) concern with a significant problem, area, or issue in the field of study;
e) promise of a significant contribution to the subject under study or problem undertaken;
f) practice of skills or application of research result which deepens or extends the applicant's professional capabilities.

The overall goal of my sabbatical project is to continue my program of scholarship in the study of type-1 diabetes through the lens of mathematics, leading to publishable results that will inspire future research with students. The research focuses on the analysis of glucose data from type-1 diabetic patients, in order to create numerical tools to help patients improve their control of blood glucose.

Over the past ten years, continuous glucose monitors (CGMs) have become readily available for use by type-1 diabetic patients, and insurance companies are beginning to accept them as a legitimate part of treatment. As the use of CGMs increases, both physicians and patients are faced with the question of how to make sense of the tidal wave of data generated by these devices. Unlike the measuring of blood sugar four to six times a day through a finger-prick or the determination four times a year of A1C (a long-term measure of average blood glucose), current CGMs generate a number every five minutes, leading to nearly 9000 readings and approximately thirty graphs of data per month, if the patient wears the monitor all the time. What can be done to mine this data and learn new facts about a patient's blood glucose management?

Prior work by the proposer on this question will be discussed in section #3. For this proposal, I plan to continue this program of scholarship, making use of a data set of blood glucose readings from diabetic patients. The following research questions will be considered:

1. Can the dynamics of an individual diabetic patient's blood glucose numbers be used to predict future short-term blood glucose behavior? (For example, is there a set of patients for which two days of relatively stable data implies with high probability a highly variable third day? If so, what counseling can be provided to these patients to adjust their blood glucose management going into the third day?)

2. Recently, Netflix (the movie rental company) ran a public contest to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences. Successful approaches involved intricate mathematical analyses of the large data set that Netflix made public, using mathematical tools such as singular value decomposition of large matrices. A key idea was to develop measures of similarities between Netflix customers. How can these approaches be adapted to blood glucose data in order to create profiles of glucose behavior which can be matched up with diabetic patients, beyond what was done previously (and described below).

In terms of the types of goals for this project, this sabbatical work is connected to items d., e., and f above. For criterion d., type-1 diabetes now afflicts nearly three million Americans, and there are on average 40 new cases per day diagnosed in the United States. Driven by applied research such as the Juvenile Diabetes Research Foundation's (JDRF) Artificial Pancreas Project, and the growing use of CGMs, there is a growing body of peer-reviewed publications in the area of applying mathematics and engineering principles towards the care of type-1 diabetics. For criterion e., currently the key statistic that is used by endocrinologists with diabetes patients is called A1C, which is a long-term measure of average blood sugar. In the work being proposed, there is a promise of a significant contribution to blood sugar management by adding other calculated statistics, based on CGM data. For criterion f, I anticipate learning new mathematical and statistical tools for the analysis of large data sets, such as de-anonymization algorithms, prediction-correction methods for time series, and methods to study data sets with "long-tailed" distributions.

The analysis that I am planning will be applied to a new data set of CGM data which includes year-long glucose time series for 451 actual type-1 diabetics. JDRF has made this data set publicly available. (See attached e-mail.)

3. Evidence of preparation.
Describe academic preparation that contributes to feasibility of the project, placing the planned activity in the context of your field. There shall be a thorough attempt by the proposer to search the literature and to place the planned activity in the context of that literature. This may take the form of literature citations, consultations, and indication of previous work in the field. If a book is being written, append an outline or table of contents to demonstrate that groundwork has been laid.

The main focus of my scholarship throughout my career has been applied mathematics. As indicated on my vita, I have pursued numerous projects, individually or in partnership with others (especially students), that feature the application of mathematics to another discipline. (Examples include the way an outfielder catches a baseball, automatically reading the Internet-based puzzle known as CAPTCHAs, and imitating the artistic style of Chuck Close through a computer algorithm.) Most of these projects have made use of the mathematical tool known as *wavelets*, and much of the work was funded in part by research experiences for undergraduates (REU) grants from the National Science Foundation. Wavelets, a mathematical tool that came to prominence in the early 1990's, are good at detecting hidden trends in data, a process which has become a very important area of intellectual study and research, in areas as diverse as genomics, sports, economics, and art history.

In 2009, I worked with two REU students to answer the following question: can wavelets be used to create "digital signatures" of type-1 diabetic patients from their blood glucose data, and with these signatures, can the patients then be "clustered" into groups with similar signatures? Through our analysis, my students and I developed a measure of blood glucose variability called the "PLA index", where PLA stands for "piecewise linear approximation". Using data from 106 actual patients, we created three clusters of patients, based on the value of the index for each person. My students presented these results at a national conference called MathFest and won one of two Janet L. Anderson Prizes for Undergraduate Research in Mathematical or Computational Biology. A manuscript has been submitted to the journal *Algorithms*.

In preparing for the 2009 program, I conducted an extensive literature search on connections between mathematics and type 1 diabetes. This included reading papers 1 to 12 in the Bibliography at the end of this section. These papers feature mathematical models of glucose in the blood (such as differential equation models), ideas about algorithms that would make sense in an artificial pancreas, and ways in which wavelets have been used to analyze data. This list of references suggests that there is great interest in the type of problems posed in section #2 above. Only recently, though, have CGMs been available for diabetic patients, so we are just beginning to have available a significant collection of actual data to analyze. There is very little published where CGM data has been analyzed beyond basic statistical analysis.

I expect for this project to branch out beyond wavelets to use other mathematical and statistical tools. During Summer 2010, I have been reading papers 13 through 16 in the Bibliography, to help me start to understand these tools. Reading papers of this type can be time-consuming, as they are often written tersely with unusual terminology or complicated formulas for which time is needed to make sense of. Sometimes, it is useful to write some computer code to simulate the mathematics in these papers and get a feel through examples of what the authors have accomplished. These papers in particular combine concepts from multiple disciplines, so at times I have had to better educate myself on the workings of capital markets, the construction of databases, and even the details of the job of quarterback in the National Football League. Looking ahead, I plan on learning about two topics from statistics that may be useful to meet my goals – "normal probability plots" and "Ripley K's measure of spatial clustering" – and I anticipate finding more articles to read between the submission of this proposal and the beginning of my sabbatical in Fall 2011.

Bibliography for this section:

1. Mathematical Models and Software Tools for the Glucose-Insulin Regulatory System and Diabetes: An Overview, by Athena Makroglou, Jiaxu Li, and Yang Kuang. *Applied Numerical Mathematics*, Vol. 56 (2006), p. 559-573.

2. A Review of Closed-Loop Algorithms for Glycemic Control in the Treatment of Type 1 Diabetes, by Joseph El Youssef, Jessica Castle and W. Kenneth Ward. *Algorithms*, Vol .2 (2009), p. 518-532.

3. Identification of Two Time-Series Models of Type 1 Diabetes Using Patient's Clinical Database, by A. Karim El-Jabali and S. L. Alousi. *Pakistan Journal of Applied Science*, Vol. 3 (2003), p. 274-279.

4. Continuous Drug Infusion for Diabetes Therapy: A Closed-Loop Control System Design, by Jiming Chen, Kejie Cao, Youxian Sun, Yang Xiao, and Xu (Kevin) Su. *EURASIP Journal on Wireless Communications and Networking*, Vol. 2008 (2008), 10 pages (online).

5. Closed-loop Insulin Delivery – The Path to Physiological Glucose Control, by G.M. Steil, A.E. Panteleon, and K. Rebrin. *Advanced Drug Delivery Reviews*, Vol. 56 (2004), p. 125-144.

6. Artificial Pancreas May Soon Be a Reality, by M. J. Friedrich. *Journal of the American Medical Association*, Vol. 301 (2009), p. 1525-1527.

7. Evaluation of a New Measure of Blood Glucose Variability in Diabetes, by Boris P. Kovatchev, *et. al. Diabetes Care*, Vol. 29 (2006), p. 2433-2438.

8. Evaluation of the Effect of Gain on the Meal Response of an Automated Closed-Loop Insulin Delivery System, by Antonios E. Panteleon, Mikhail Loutseiko, Garry M. Steil, and Kerstin Rebrin. *Diabetes*, Vol. 55 (2006), p. 1995-2000.

9. A Wavelet-Based Multiresolution PID Controller, by Shahid Parvez and Zhiqiang Gao. *IEEE Transactions on Industry Applications,* Vol. 41 (2005), p. 537-543.

10. Wavelets in Time-Series Analysis, by Guy P. Nason and Rainer von Sachs. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, Vol. 357 (1999), p. 2511-2526.

11. A Survey on Wavelet Applications in Data Mining, by Tao Li, *et. al. ACM SIGKDD Explorations Newsletter*, Vol. 4 (2002), p. 49-68.

12. Laplacian eigenmaps for dimensionality reduction and data representation, by M. Belkin and P. Niyoki. *Neural Computation*, Vol. 15 (2003), p. 373-396.

13. Robust De-anonymization of Large Datasets (How to Break Anonymity of the Netflix Prize Dataset), by Arvind Narayanan and Vitaly Shmatikov. Pre-print posted to the arXiv (2008).

14. Leverage Causes Fat Tails and Clustered Volatility, by Stefan Thurner, J. Doyne Farmer and John Geanakoplos. Pre-print posted to the arXiv (2010).

15. On the Use of Gromov-Hausdorff Distances for Shape Comparisons, by F. Memoli. *Proceedings of the Symposium on Point Based Graphics* (2007).

16. Untangling Skill and Luck, by Michael Mauboussin. Posted online at Legg Mason.

4. Project Plan.

Describe the sabbatical project. Show how that plan relates to its goals and objectives.

Come August 2011, I will have done enough intense, detailed reading to have developed very specific and answerable questions about the CGM data that I have access to, in order to make progress on the two research questions that I have proposed. For instance, for research question #1, a good question to start with is the following: what is the statistical distribution of how long a type-1 diabetic's blood sugar stays "in range" (between 70 and 150 mg/dL)? This would lead to identifying which of the patients in the database appear to be the most skilled at managing their blood sugar, and my understanding of the data would deepen. If we view the management of blood sugar as a performance, there are results from other areas of performance analysis (e.g. managing stock portfolios, sports) that suggest that performance can be divided into that resulting from skill and that resulting from luck. By identifying the characteristics of "skill" in diabetes management, predictions can be made. It is much more difficult to make predictions if luck predominates. (Consider a coin flip, or even a sequence of coin flips.) So, the next activity would be to use statistical techniques to divide the contributions of skill and luck.

For research question #2, the key issue will be to define a measure of similarity among diabetics for which application of the Netflix algorithm, or some variation, would lead to useful results. I have some preliminary results in this area from the 2009 REU program, but an important activity here would be to write computer code to model the Netflix algorithm, while developing criteria for "useful results". The latter would come from my reading of the literature. After creating the code, which will not be a simple task, different measures of similarity will be tested, using the CGM data.

I expect the approaches just described will lead to the outcome of answering questions #1 and #2, meetings goals d, e, and f in Section 2. I believe I can make significant progress on both questions during Fall 2011, in the sense that both questions have a time-limited piece and an open-ended piece. For example, with question #1, the time-limited piece is answering specific questions about predictions, such as the one about whether or not the third day is stable or highly variable. However, the answer to the broad question is probably "Yes", but the implied question is "how can predictions be made?", and I expect to provide one answer during Fall 2011. For question #2, I will definitely be able to implement the Netflix algorithm and apply it to the CGM data that I have, but I cannot predict what the results will be. I believe, based on the REU 2009 work, that there is some structure to the set of type-1 diabetic patients, so that some interesting results will arise from the Netflix algorithm.

Much of the work will involve coding to either clean up the CGM data or analyze it via algorithms of my design. So, I expect to use a significant amount of time creating and coding algorithms based on my reading and some pencil-and-paper mathematical analysis. I will spend time thinking about the results that I get from computer calculations, and what they mean.

5. Timetable.

Indicate estimated dates for each of the significant steps in the project plan. Be as specific as possible. Include an explanation showing whether the project can be completed in the time available. If the sabbatical leave is being used to begin a longer term project, state when you expect the whole activity to be completed.*

The first part of my plan was completed recently, as I acquired from JDRF a new data set of CGM data which includes year-long glucose time series for 451 actual type-1 diabetics.

May 2011: Further literature search and reading, along with preparation to mentor two students at the 2011 REU. I anticipate a research project next summer that will be a preliminary step towards answering the questions posed in Section #2 above, such as converting a blood glucose data set into a sparse matrix and analyzing the matrix through tools from linear algebra. Any project designed for REU students would be at levels below what I will be considering during my sabbatical, but may lead to a mathematical tool that I can use during my sabbatical research.

June-August 2011: The REU program.

Mid-August 2011: Sabbatical begins. I expect one month of further reading and learning before turning my attention to analyzing actual blood glucose data.

September-October 2011: Tackle question #1 from Section #2 above. Because I have some strong insights of how to proceed with this problem, I will explore this question first. I anticipate having results by the end of October.

November-December 2011: Tackle question #2.

Sometime during Fall 2011 semester: Attend a national conference related to data mining. Because of the popularity of this topic, there are many conferences available in a given year to attend. (For example, the ACM-Knowledge Discovery and Data Mining Conference in San Diego, August 2011, or the SAS Data Mining Conference, October 2011.)

During the sabbatical: Work with my REU students to finalize a manuscript based on our work, with the goal of submitting the manuscript by the end of 2011.

By April 2012: Decide on one or two conferences to attend and present my work. A few ideas include the Michigan section of the Mathematical Association of America conference in May 2012, the Society for Industrial and Applied Mathematicians annual meeting in July 2012, or the Joint Mathematics Meetings in January 2013.

May-June 2012: Due to my work as Unit Head, I don't anticipate time during Winter 2012 to write up my results from my sabbatical. I will do so at this time.

6. Benefit to one's own or other units.
A clear relationship between the proposed sabbatical leave and a proposer's academic unit shall be demonstrated. If your project is to benefit a unit other than your home unit, describe that situation. Attach signed, written verification of that benefit from the head of that other unit.

The Department of Mathematics values clear, on-going programs of scholarship among its members, and this sabbatical will enhance my research program and provide a foundation for work by myself and in collaboration with students and/or other faculty the next several years. (My last sabbatical in 2000 did just that, where I expanded my knowledge of wavelets, Fourier analysis, and applied mathematics that led to several successful undergraduate research projects and a textbook on advanced engineering mathematics.)

7. Arrangements with people or other institutions.
If the project requires collaboration with other institutions or persons, describe the collaboration and provide evidence that the institutions or persons involved agree with the arrangements.** If not, please indicate that no collaboration is necessary or planned. For books, indicate interest of potential publisher if available. You may append correspondence.

In conjunction with this proposal, I am applying to the Juvenile Diabetes Research Foundation (JDRF) for a small grant to support this research. However, the grant is not necessary for my sabbatical.

8. Other relevant information.

None.

9. Appendices.

a. Attach copies of previous sabbatical reports in the Appendix.

b. Attach a copy of your curriculum vitae in the appendix.
This should be current and complete. Include a list of publications, exhibitions, performances, and productions.

c. Attach letters of support or other supporting documentation.

D. If modification is made after a step in the review process but prior to the College Personnel Committee recommendation, incorporate the modification into the appropriate section of the final proposal as opposed to appending pages. If changes are proposed after the final approval but before or during the sabbatical period, prior consultation with the chair of the unit should occur and a new formal plan document must be approved by the Dean and the Provost before implementation of the proposed plan.

E. Final Report. Each faculty member returning from sabbatical leave will provide a written account of the sabbatical activities and accomplishments and deposit copies with his/her appointing officer and unit head, the Provost, the President, the Research and Development Center, and the library. The report shall be filed no later than the end of the first semester after return to campus and shall include an account of the financial remuneration received during the sabbatical leave. Failure to complete the approved sabbatical project, submit a final report or participate in dissemination of sabbatical work may negatively affect an annual salary adjustment and decisions on future sabbatical proposals.

---

*In such a case, your sabbatical report is to be made as scheduled by the policy stated in the Administrative Manual even if the entire project is unfinished. Report on whatever was accomplished during the sabbatical leave.

**If there is a possibility of creating Intellectual Property during a sabbatical and you are involved in external collaboration with another university, business or organization, it is important to review the GVSU and collaborating institution Intellectual Property Policy (IPP) prior to initiating work activities. For support in the IPP review contact GVSU legal counsel.

Revised 6/06