Stem Cell Reports
Commentary
OPEN ACCESS

Lessons Learned from Pioneering Neural Stem Cell Studies

Sally Temple[1,*] and Lorenz Studer[2]
[1]Neural Stem Cell Institute, Rensselaer, NY 12144, USA
[2]Memorial Sloan Kettering Cancer Center, Center for Stem Cell Biology, New York, NY 10065, USA
*Correspondence: sallytemple@neuralsci.org
http://dx.doi.org/10.1016/j.stemcr.2017.01.024

As stem cell products are increasingly entering early stage clinical trials, we are learning from experience about how cell products may be best assessed for safety and efficacy. In two papers published in this issue of *Stem Cell Reports*, a human neural stem cell product, HuCNS-SC, failed to demonstrate efficacy in central nervous system repair in two different animal models (Anderson et al., 2017; Marsh et al., 2017), although closely related research-grade cell products showed evidence of efficacy. This indicates the need for increased cell characterization to determine comparability of lots proposed for pre-clinical and clinical use. Without such improvements, pre-clinical data supporting a clinical study might not adequately reflect the performance of subsequent batches of cells intended for use in patients.

In this issue of *Stem Cell Reports*, two critically important studies describe negative results of a neural stem cell (NSC) product (HuCNS-SC) intended for clinical use in a model of cervical spinal cord injury (SCI) (Anderson et al., 2017) and in a model of Alzheimer's disease (Marsh et al., 2017). Anderson et al. reported that they relayed their negative results to the company 6 months ahead of the first patient dosing, and yet the decision was made to continue with a cervical SCI clinical trial. Data obtained from the first six patients in this clinical Pathway Study showed an initial small improvement that did not persist at later study time points (up to 1 year), and a decision was made to terminate the trial in May 2016; for business reasons, the company providing HuCNS-SC, StemCells Inc., folded. The two reports raise several important questions. Why did research grade NSCs show benefit in pre-clinical models of cervical SCI whereas a comparable clinical lot did not (Anderson et al., 2017)? Was the preclinical failure predictive of failure for the clinical Pathway Study? And how should stakeholders—regulatory officials, physicians, and participants—be best informed about failed efficacy data in order to decide whether to continue with or participate in a clinical study? The need for discussion about how cell products are characterized and tested for comparability and how these data are used is heightened by the drive to accelerate the approval process for regenerative therapy products, already accomplished in several countries and expected to result from the US 21st Century Cures Act.

After demonstrating efficacy of research-grade HuCNS-SC cells in murine thoracic spinal cord injury models, the Cummings lab was excited to explore the application of this product to the more severe cervical injury. Anderson et al. (2017) performed a controlled, masked study to assess the efficacy of HuCNS-SC for cervical SCI using a clinical cell line (CCL) supplied by StemCells Inc. A "comparable" research grade cell line (RCL) was also provided by StemCells Inc. All the cell preparations were shipped overnight with appropriate monitoring and transplanted on day of receipt. The RCL product showed efficacy for SCI in immunodeficient Rag2γ mice injected with 75,000 cells at 9 days or 60 days post injury. Locomotor function was significantly improved at 12 weeks when RCL NSCs were transplanted at 9 days post injury, with less effect for 60 day post-injury transplants. The CCL groups, however, showed no locomotor improvement at either time point and, in fact, a possible worsening of outcomes associated with more extensive CCL engraftment. Based on the lack of efficacy in the CCL studies, these results might explain the lack of efficacy in the Pathway Study.

In a companion study aimed at demonstrating the therapeutic potential of StemCells Inc.'s HuCNS-SC in an Alzheimer's disease animal model, clinical-grade cells were transplanted into the brain of Rag-5xfAD mice. Despite robust engraftment, treated animals did not improve cognition, increase BDNF, or increase synaptic density at 5 months after transplantation. This was in contrast to prior studies using a research grade HuCNS-SC preparation provided by StemCells Inc. that showed promising results in an Alzheimer's disease model at 1 month post transplantation (Ager et al., 2015). In addition, the longer duration study found periventricular cell clusters in a subset of animals—clusters resembling rare neurocytoma tumors according to one of the pathologists. This study amplifies concern about differences between the test cell preparations and points to the importance of performing longer-term functional and safety studies in pre-clinical models of central nervous system repair.

What may explain the differences in performance between manufactured cell lots? Typically, a research-grade cell product is first tested in animals to show positive effects. Subsequently, the manufacturing process is brought to a clinical level using current good manufacturing practice (cGMP) designed to produce a reliably consistent product through carefully documented characterization of the

source cells, components used in the manufacturing process, and tests of the final cell product (FCP) for identity, purity, and potency. We presume that StemCells Inc. released the research and clinical grade cells for use in these efficacy studies after performing appropriate testing and comparability studies, as the intention was to gain supportive evidence for clinical studies. Assuming no material deviations from the manufacturing protocol occurred that could explain the difference in cell behavior between the products tested, the assumption that specific manufacturing and release criteria produce comparable FCPs that perform similarly is now challenged.

The StemCells Inc. manufacturing process is proprietary, and we do not know exactly how HuCNS-SC was made and tested to determine comparability. We do know that HuCNS-SC is derived from fetal tissue and, hence, each new source is from a different donor with unique genetic makeup. Furthermore, NSCs vary significantly in types and numbers of neurons, astrocytes, and oligodendrocytes produced, depending on the origin location in the nervous system and the developmental stage at which they are extracted, so heterogeneity in the starting tissue is a challenge. Variations in growth and differentiation reagents can lead to further differences in FCP composition. Indeed, for the RCL and CCL preparations sent from StemCells Inc. on ice overnight and treated the same way at the testing lab, the CCL had more cell debris and lower viability than the RCL. These obvious differences signaled concerns which were relayed to StemCells Inc., but the preclinical animal studies continued, therefore presumably the measurements were within the bounds of acceptance criteria for FCP release.

Cells are highly complex systems that are dynamic and responsive to the environment. The high cost of testing numerous markers to define cell products in the cGMP setting, however, often leads to a "defining minimum" employing the fewest markers necessary to identify desired cells and contaminants and as a surrogate for potency. Of the defining characteristics, the potency of a cell product is the most challenging to ascertain. Potency implies that the mechanism of action is understood, which is difficult for a cell product, given its multimodal actions. The FDA understands this challenge and does not require a definitive potency assay for early stage clinical trial. The absence of potency criteria to define different batches of clinical grade cells, however, leaves room for variation in FCP behavior.

How might the product evaluation process be improved? Recent advances enable us to define cell products more comprehensively via whole genome sequencing, transcriptomics, epigenomics, and proteomics at the population level and with single cell omics to determine population heterogeneity. In order to implement these methods in the cell production process, however, equipment and protocols need to be qualified for GMP, which is costly, as are the tests themselves. Furthermore, the rigorous studies that would be needed to correlate multiple markers with in vivo efficacy outcomes will further increase development expense and time to enter the clinic. Still, the failure to advance cell product characterization means that we continue to run the risk of failures such as those exemplified by these two papers and the termination of the Pathway Study. A goal for the field is to strike a balance that allows practical, yet enhanced, characterization that recognizes meaningful differences in cellular products.

With the increasing use of pluripotent-derived cell types for CNS repair, it has become possible to generate sufficient FCP to allow both pre-clinical testing and clinical study on the same lot. This resolves uncertainty regarding FCP potency and provides a fully tested "off-the-shelf" product, albeit one limited by having to directly transplant cryopreserved product and to start with a single stem cell source capable of generating lots of adequate size. Furthermore, the approach does not resolve the longer-term problem, as eventually even a large lot of FCP will run out. Moreover, in the case of individualized iPSC-based treatment or the use of HLA-matched iPSC banks, it may not be feasible to test every FCP in definitive pre-clinical animal studies. Given the possibility for variation demonstrated by the papers in this issue, should each new cell line be given a unique identifier that is disclosed to investigators and participants in the clinical trial so that they understand which cells are provided and the levels of functional characterization performed on those cells? When the cell product is labeled the same (in this case all are labeled HuCNS-SC), how can patients and physicians know the extent of testing that has been performed on a particular line and understand the risks contributed by product variability in order to make an informed decision on whether to participate in a trial? This point is discussed along with further background to their studies on StemCells Inc.'s products and implications for clinical advancement of cell therapies (Anderson and Cummings, 2016).

The ISSCR has recently released updated guidelines on stem cell research and clinical translation (Daley et al., 2016; ISSCR, 2016) that stress the importance of rigorous preclinical testing and recommends the publication of preclinical studies in full to allow assessment of the strength of the evidence supporting human translation. In addition, given the highly specialized knowledge required to assess cell characteristics and in vivo efficacy tests, we suggest that peer review by experts serving in an advisory function to regulatory authorities would help ensure that cells enter clinical trials based on sound scientific rationale with

robust manufacturing and animal efficacy data in addition to a rigorous safety package to support clinical trial FDA allowance. Confidential external peer review would ensure that higher quality trials are supported, which is important to direct limited resources to studies that have the greatest potential for success.

In conclusion, these two papers highlight the need to improve cell product characterization and to establish potency assays that correlate with efficacy, which will provide more confidence in the comparability of different manufactured lots. More emphasis on robust and reliable potency assays, as called for in the ISSCR guidelines, would push the cell therapy field to overall higher standards. Without this, regulators may require sponsors to perform efficacy tests on every clinical lot, which is a costly solution. On the other hand, we also must recognize the limitations of many animal models, and with this, the considerable challenge of defining mechanism of action for a complex, cell-based therapeutic in what are poor approximations of the human disease. Improvements in the accuracy of animal models should be a priority, but ultimately, we do need to get into the clinic and perform the tests in patients, recognizing that the difficulty in extrapolating from animal to human will have an inherent failure rate. We share the excitement at the entrance of stem-cell based regenerative therapies to the clinical repertoire, and recognizing the urgent need for many disease indications, timely improvements in the cell manufacturing and testing process are required, because the health and safety of patients must remain our uppermost concern.

ACKNOWLEDGMENTS

We are grateful to Jeff Stern and Heather Rooke for their valuable comments on the manuscript.

REFERENCES

Ager, R.R., Davis, J.L., Agazaryan, A., Benavente, F., Poon, W.W., LaFerla, F.M., and Blurton-Jones, M. (2015). Human neural stem cells improve cognition and promote synaptic growth in two complementary transgenic models of Alzheimer's disease and neuronal loss. Hippocampus *25*, 813–826.

Anderson, A.J., and Cummings, B.J. (2016). Achieving informed consent for cellular therapies: a preclinical translational research perspective on regulations versus a dose of reality. J. Law Med. Ethics *44*, 394–401.

Anderson, A.J., Piltti, K.M., Hooshmand, M.J., Nishi, R.A., and Cummings, B.J. (2017). Preclinical efficacy failure of human CNS-derived stem cells for use in the pathway study of cervical spinal cord injury. Stem Cell Reports *8*, this issue, 249–263.

Daley, G.Q., Hyun, I., Apperley, J.F., Barker, R.A., Benvenisty, N., Bredenoord, A.L., Breuer, C.K., Caulfield, T., Cedars, M.I., Frey-Vasconcells, J., et al. (2016). Setting global standards for stem cell research and clinical translation: the 2016 ISSCR guidelines. Stem Cell Reports *6*, 787–797.

International Society for Stem Cell Research (2016). Guidelines for Stem Cell Research and Translation. http://www.isscr.org/guidelines2016.

Marsh, S.E., Yeung, S.T., Torres, M., Lau, L., Davis, J.L., Monuki, E.S., Poon, W.W., and Blurton-Jones, M. (2017). HuCNS-SC human NSCs fail to differentiate, form ectopic clusters, and provide no cognitive benefits in a transgenic model of Alzheimer's disease. Stem Cell Reports *8*, this issue, 235–248.